(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 August 2001 (30.08.2001)

PCT

(10) International Publication Number
WO 01/62922 A2

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, 16/18, A61K 38/17, A01K 67/027, G01N 33/50, 33/68, C12Q 1/68

(21) International Application Number: PCT/US01/05896

(22) International Filing Date: 21 February 2001 (21.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/185,213	24 February 2000 (24.02.2000)	US
60/205,232	16 May 2000 (16.05.2000)	US
60/205,285	17 May 2000 (17.05.2000)	US
60/205,323	17 May 2000 (17.05.2000)	US
60/205,287	17 May 2000 (17.05.2000)	US
60/205,324	17 May 2000 (17.05.2000)	US
60/205,286	17 May 2000 (17.05.2000)	US

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **PANZER, Scott, R.** [US/US]; 571 Bobolink Circle, Sunnyvale, CA 94087 (US). **SPIRO, Peter, A.** [US/US]; 410 Sheridan Avenue #333, Palo Alto, CA 94306 (US). **BANVILLE, Steven, C.** [US/US]; 365 Monroe Drive, Palo Alto, CA 94306 (US). **SHAH, Purvi** [IN/US]; 859 Salt Lake Drive, San Jose, CA 95133 (US). **CHALUP, Michael, S.** [US/US]; 183 Acalanes Drive, Apt. #6, Sunnyvale, CA 94086 (US). **CHANG, Simon, C.** [US/US]; 1901 Rock Street, #103, Mountain View, CA 94043 (US). **CHEN, Alice** [US/US]; 4405 Norwalk Drive, #22, San Jose, CA 95129 (US). **D'SA, Steven, A.** [CA/US]; Apartment #13, 40 Newell Road, East Palo Alto, CA 94303 (US). **AMSHEY, Stefan** [US/US]; 1605 20th Street, San Francisco, CA 94107 (US). **DAHL, Christopher, R.** [US/US]; 41277 Roberts Avenue, #6, Fremont, CA 94538 (US). **DAM, Tam, C.** [US/US]; 2180 Mendota Way, San Jose, CA (US). **DANIELS, Susan, E.** [GB/US]; 136 Seale Avenue, Palo Alto, CA 94301 (US). **DUFOUR, Gerard, E.** [US/US]; 5327 Greenridge Road, Castro Valley, CA 94552-2619 (US). **FLORES, Vincent** [US/US]; 35000 Begonia Street, Union City, CA 94587 (US). **FONG, Willy, T.** [US/US]; 572 Cambridge Street, San Francisco, CA 94134 (US). **GREENAWALT, Lila, B.** [US/US]; 1596 Ballantree Way, San Jose, CA 95118-2106 (US). **HILLMAN, Jennifer, L.** [US/US]; 230 Monroe Drive, #17, Mountain View, CA 94040 (US). **JONES, Anissa, L.** [US/US]; 445 South 15th Street, San Jose, CA 95112 (US). **LIU, Tommy, F.** [US/US]; 201 Ottilia Street, Daly City, CA 94014 (US). **ROSEBERRY, Ann, M.** [US/US]; 725 Sapphire Street, Redwood City, CA 94061 (US). **ROSEN, Bruce, H.** [US/US]; 177 Hanna Way, Menlo Park, CA 94025 (US). **RUSSO, Frank, D.** [US/US]; 939 Rosette Court, Sunnyvale, CA 94086 (US). **STOCKDREHER, Theresa, K.** [US/US]; 1596 Ontario Drive, #2, Sunnyvale, CA 94087 (US). **DAFFO, Abel** [US/US]; 1750 Stokes Street, #70, San Jose, CA 95126 (US). **WRIGHT, Rachel, J.** [NZ/US]; 333 Anna Avenue, Mountain View, CA 94043 (US). **YAP, Pierre, E.** [US/US]; 201 Happy Hollow Court, Lafayette, CA 94549-6243 (US). **YU, Jimmy, Y.** [US/US]; 3655 Wyndham Drive, Fremont, CA 94536 (US). **BRADLEY, Diana, L.** [US/US]; 3260 Crystal Heights Drive, Soquel, CA 95073 (US). **BRATCHER, Shawn, R.** [US/US]; 550 Ortega Avenue, #B321, Mountain View, CA 94040 (US). **CHEN, Wensheng** [CN/US]; 210 Easy Street #25, Mountain View, CA 94043 (US). **COHEN, Howard, J.** [US/US]; 3272 Cowper Street, Palo Alto, CA 94306-3004 (US). **HODGSON, David, M.** [US/US]; 2795 Windwood Drive, Apartment 165, Ann Arbor, MI 48105 (US). **LINCOLN, Stephen, E.** [US/US]; 725 Sapphire Street, Redwood City, CA 94061 (US). **JACKSON, Stuart** [US/US]; 230 Monroe Drive, #17, Mountain View, CA 94040 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian

[Continued on next page]

(54) Title: MOLECULES FOR DISEASE DETECTION AND TREATMENT

(57) Abstract: The present invention provides purified disease detection and treatment molecule polynucleotides (mddt). Also encompassed are the polypeptides (MDDT) encoded by mddt. The invention also provides for the use of mddt, or complements, oligonucleotides, or fragments thereof in diagnostic assays. The invention further provides for vectors and host cells containing mddt for the expression of MDDT. The invention additionally provides for the use of isolated and purified MDDT to induce anitbodies and to screen libraries of compounds and the use of anti-MDDT antibodies in diagnostic assays. Also provided are microarrays containing mddt and methods of use.

WO 01/62922 A2

patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

MOLECULES FOR DISEASE DETECTION AND TREATMENT

TECHNICAL FIELD

The present invention relates to molecules for disease detection and treatment and to the use of these sequences in the diagnosis, study, prevention, and treatment of diseases associated with, as well as effects of exogenous compounds on, the expression of molecules for disease detection and treatment.

BACKGROUND OF THE INVENTION

The human genome is comprised of thousands of genes, many encoding gene products that function in the maintenance and growth of the various cells and tissues in the body. Aberrant expression or mutations in these genes and their products is the cause of, or is associated with, a variety of human diseases such as cancer and other cell proliferative disorders. The identification of these genes and their products is the basis of an ever-expanding effort to find markers for early detection of diseases, and targets for their prevention and treatment.

For example, cancer represents a type of cell proliferative disorder that affects nearly every tissue in the body. A wide variety of molecules, either aberrantly expressed or mutated, can be the cause of, or involved with, various cancers because tissue growth involves complex and ordered patterns of cell proliferation, cell differentiation, and apoptosis. Cell proliferation must be regulated to maintain both the number of cells and their spatial organization. This regulation depends upon the appropriate expression of proteins which control cell cycle progression in response to extracellular signals such as growth factors and other mitogens, and intracellular cues such as DNA damage or nutrient starvation. Molecules which directly or indirectly modulate cell cycle progression fall into several categories, including growth factors and their receptors, second messenger and signal transduction proteins, oncogene products, tumor-suppressor proteins, and mitosis-promoting factors. Aberrant expression or mutations in any of these gene products can result in cell proliferative disorders such as cancer. Oncogenes are genes generally derived from normal genes that, through abnormal expression or mutation, can effect the transformation of a normal cell to a malignant one (oncogenesis). Oncoproteins, encoded by oncogenes, can affect cell proliferation in a variety of ways and include growth factors, growth factor receptors, intracellular signal transducers, nuclear transcription factors, and cell-cycle control proteins. In contrast, tumor-suppressor genes are involved in inhibiting cell proliferation. Mutations which cause reduced or loss of function in tumor-suppressor genes result in aberrant cell proliferation and cancer. Thus a wide variety of genes and their products have been found that are associated with cell proliferative disorders such as cancer, but many more may exist that are yet to be discovered.

DNA-based arrays can provide a simple way to explore the expression of a single polymorphic

gene or a large number of genes. When the expression of a single gene is explored, DNA-based arrays are employed to detect the expression of specific gene variants. For example, a p53 tumor suppressor gene array is used to determine whether individuals are carrying mutations that predispose them to cancer. A cytochrome p450 gene array is useful to determine whether individuals have one of a number of specific mutations that could result in increased drug metabolism, drug resistance or drug toxicity.

DNA-based array technology is especially relevant for the rapid screening of expression of a large number of genes. There is a growing awareness that gene expression is affected in a global fashion. A genetic predisposition, disease or therapeutic treatment may affect, directly or indirectly, the expression of a large number of genes. In some cases the interactions may be expected, such as when the genes are part of the same signaling pathway. In other cases, such as when the genes participate in separate signaling pathways, the interactions may be totally unexpected. Therefore, DNA-based arrays can be used to investigate how genetic predisposition, disease, or therapeutic treatment affects the expression of a large number of genes.

The discovery of new molecules for disease detection and treatment satisfies a need in the art by providing new compositions which are useful in the diagnosis, study, prevention, and treatment of diseases associated with, as well as effects of exogenous compounds on, the expression of molecules for disease detection and treatment.

SUMMARY OF THE INVENTION

The present invention relates to human disease detection and treatment molecule polynucleotides (*mddt*) as presented in the Sequence Listing. The *mddt* uniquely identify genes encoding structural, functional, and regulatory disease detection and treatment molecules.

The invention provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; c) a polynucleotide sequence complementary to a); d) a polynucleotide sequence complementary to b); and e) an RNA equivalent of a) through d). In one alternative, the polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45. In another alternative, the polynucleotide comprises at least 60 contiguous nucleotides of a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; c) a polynucleotide sequence complementary to a); d) a

polynucleotide sequence complementary to b); and e) an RNA equivalent of a) through d). The invention further provides a composition for the detection of expression of disease detection and treatment molecule polynucleotides comprising at least one isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; c) a polynucleotide sequence complementary to a); d) a polynucleotide sequence complementary to b); and e) an RNA equivalent of a) through d); and a detectable label.

The invention also provides a method for detecting a target polynucleotide in a sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; c) a polynucleotide sequence complementary to a); d) a polynucleotide sequence complementary to b); and e) an RNA equivalent of a) through d). The method comprises a) amplifying said target polynucleotide or a fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention also provides a method for detecting a target polynucleotide in a sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; c) a polynucleotide sequence complementary to a); d) a polynucleotide sequence complementary to b); and e) an RNA equivalent of a) through d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof. In one alternative, the probe comprises at least 30 contiguous nucleotides. In another alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a recombinant polynucleotide comprising a promoter sequence operably linked to an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a

polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; c) a polynucleotide sequence complementary to a); d) a polynucleotide sequence complementary to b); and e) an RNA equivalent of a) through d). In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide. In a further alternative, the invention provides a method for producing a disease detection and treatment molecule polypeptide, the method comprising a) culturing a cell under conditions suitable for expression of the disease detection and treatment molecule polypeptide, wherein said cell is transformed with the recombinant polynucleotide, and b) recovering the disease detection and treatment molecule polypeptide so expressed.

The invention also provides a purified disease detection and treatment molecule polypeptide (MDDT) encoded by at least one polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45. Additionally, the invention provides an isolated antibody which specifically binds to the disease detection and treatment molecule polypeptide. The invention further provides a method of identifying a test compound which specifically binds to the disease detection and treatment molecule polypeptide, the method comprising the steps of a) providing a test compound; b) combining the disease detection and treatment molecule polypeptide with the test compound for a sufficient time and under suitable conditions for binding; and c) detecting binding of the disease detection and treatment molecule polypeptide to the test compound, thereby identifying the test compound which specifically binds the disease detection and treatment molecule polypeptide.

The invention further provides a microarray wherein at least one element of the microarray is an isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; c) a polynucleotide sequence complementary to a); d) a polynucleotide sequence complementary to b); and e) an RNA equivalent of a) through d). The invention also provides a method for generating a transcript image of a sample which contains polynucleotides. The method comprises a) labeling the polynucleotides of the sample, b) contacting the elements of the microarray with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and c) quantifying the expression of the polynucleotides in the sample.

Additionally, the invention provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; b) a naturally occurring polynucleotide sequence having

at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; c) a polynucleotide sequence complementary to a); d) a polynucleotide sequence complementary to b); and e) an RNA equivalent of a) through d). The method comprises a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; ii) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; ii) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45; iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv), and alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

The invention further provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:46-90, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:46-90, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:46-90, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:46-90. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:46-90.

DESCRIPTION OF THE TABLES

Table 1 shows the sequence identification numbers (SEQ ID NO:s) and template identification numbers (template IDs) corresponding to the polynucleotides of the present invention, along with their GenBank hits (GI Numbers), probability scores, and functional annotations corresponding to the GenBank hits.

Table 2 shows the sequence identification numbers (SEQ ID NO:s) and template identification numbers (template IDs) corresponding to the polynucleotides of the present invention, along with polynucleotide segments of each template sequence as defined by the indicated "start" and "stop" nucleotide positions. The reading frames of the polynucleotide segments and the Pfam hits, Pfam descriptions, and E-values corresponding to the polypeptide domains encoded by the polynucleotide segments are indicated.

Table 3 shows the sequence identification numbers (SEQ ID NO:s) and template identification numbers (template IDs) corresponding to the polynucleotides of the present invention, along with polynucleotide segments of each template sequence as defined by the indicated "start" and "stop" nucleotide positions. The reading frames of the polynucleotide segments are shown, and the polypeptides encoded by the polynucleotide segments constitute either signal peptide (SP) or transmembrane (TM) domains, as indicated. The membrane topology of the encoded polypeptide sequence is indicated, the N-terminus (N) listed as being oriented to either the cytosolic (in) or non-cytosolic (out) side of the cell membrane or organelle.

Table 4 shows the sequence identification numbers (SEQ ID NO:s) corresponding to the polynucleotides of the present invention, along with component sequence identification numbers (component IDs) corresponding to each template. The component sequences, which were used to assemble the template sequences, are defined by the indicated "start" and "stop" nucleotide positions along each template.

Table 5 shows the tissue distribution profiles for the templates of the invention.

Table 6 shows the sequence identification numbers (SEQ ID NO:s) corresponding to the polypeptides of the present invention, along with the reading frames used to obtain the polypeptide segments, the lengths of the polypeptide segments, the "start" and "stop" nucleotide positions of the polynucleotide sequences used to define the encoded polypeptide segments, the GenBank hits (GI Numbers), probability scores, and functional annotations corresponding to the GenBank hits.

Table 7 summarizes the bioinformatics tools which are useful for analysis of the polynucleotides of the present invention. The first column of Table 7 lists analytical tools, programs, and algorithms, the second column provides brief descriptions thereof, the third column presents

appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences).

DETAILED DESCRIPTION OF THE INVENTION

Before the nucleic acid sequences and methods are presented, it is to be understood that this invention is not limited to the particular machines, methods, and materials described. Although particular embodiments are described, machines, methods, and materials similar or equivalent to these embodiments may be used to practice the invention. The preferred machines, methods, and materials set forth are not intended to limit the scope of the invention which is limited only by the appended claims.

The singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. All technical and scientific terms have the meanings commonly understood by one of ordinary skill in the art. All publications are incorporated by reference for the purpose of describing and disclosing the cell lines, vectors, and methodologies which are presented and which might be used in connection with the invention. Nothing in the specification is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

Definitions

As used herein, the lower case "mddt" refers to a nucleic acid sequence, while the upper case "MDDT" refers to an amino acid sequence encoded by mddt. A "full-length" mddt refers to a nucleic acid sequence containing the entire coding region of a gene endogenously expressed in human tissue.

"Adjuvants" are materials such as Freund's adjuvant, mineral gels (aluminum hydroxide), and surface active substances (lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol) which may be administered to increase a host's immunological response.

"Allele" refers to an alternative form of a nucleic acid sequence. Alleles result from a "mutation," a change or an alternative reading of the genetic code. Any given gene may have none, one, or many allelic forms. Mutations which give rise to alleles include deletions, additions, or substitutions of nucleotides. Each of these changes may occur alone, or in combination with the others, one or more times in a given nucleic acid sequence. The present invention encompasses allelic mddt.

"Amino acid sequence" refers to a peptide, a polypeptide, or a protein of either natural or synthetic origin. The amino acid sequence is not limited to the complete, endogenous amino acid

sequence and may be a fragment, epitope, variant, or derivative of a protein expressed by a nucleic acid sequence.

"Amplification" refers to the production of additional copies of a sequence and is carried out using polymerase chain reaction (PCR) technologies well known in the art.

"Antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind MDDT polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or peptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

"Antisense sequence" refers to a sequence capable of specifically hybridizing to a target sequence. The antisense sequence may include DNA, RNA, or any nucleic acid mimic or analog such as peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine.

"Antisense sequence" refers to a sequence capable of specifically hybridizing to a target sequence. The antisense sequence can be DNA, RNA, or any nucleic acid mimic or analog.

"Antisense technology" refers to any technology which relies on the specific hybridization of an antisense sequence to a target sequence.

A "bin" is a portion of computer memory space used by a computer program for storage of data, and bounded in such a manner that data stored in a bin may be retrieved by the program.

"Biologically active" refers to an amino acid sequence having a structural, regulatory, or biochemical function of a naturally occurring amino acid sequence.

"Clone joining" is a process for combining gene bins based upon the bins' containing sequence information from the same clone. The sequences may assemble into a primary gene transcript as well as one or more splice variants.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing (5'-A-G-T-3' pairs with its complement 3'-T-C-A-5').

A "component sequence" is a nucleic acid sequence selected by a computer program such as PHRED and used to assemble a consensus or template sequence from one or more component sequences.

A "consensus sequence" or "template sequence" is a nucleic acid sequence which has been assembled from overlapping sequences, using a computer program for fragment assembly such as the GELVIEW fragment assembly system (Genetics Computer Group (GCG), Madison WI) or using a relational database management system (RDMS).

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain.

"Deletion" refers to a change in either a nucleic or amino acid sequence in which at least one nucleotide or amino acid residue, respectively, is absent.

"Derivative" refers to the chemical modification of a nucleic acid sequence, such as by replacement of hydrogen by an alkyl, acyl, amino, hydroxyl, or other group.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

"E-value" refers to the statistical probability that a match between two sequences occurred by chance.

A "fragment" is a unique portion of mddt or MDDT which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 10 to 1000 contiguous amino acid residues or nucleotides. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous amino acid residues or nucleotides in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing and the figures, may be encompassed by the present embodiments.

A fragment of mddt comprises a region of unique polynucleotide sequence that specifically identifies mddt, for example, as distinct from any other sequence in the same genome. A fragment of mddt is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish mddt from related polynucleotide sequences. The precise length of a fragment of mddt and the region of mddt to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of MDDT is encoded by a fragment of mddt. A fragment of MDDT comprises a region of unique amino acid sequence that specifically identifies MDDT. For example, a fragment of MDDT is useful as an immunogenic peptide for the development of antibodies that specifically recognize MDDT. The precise length of a fragment of MDDT and the region of MDDT to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" nucleotide sequence is one containing at least a start site for translation to a protein sequence, followed by an open reading frame and a stop site, and encoding a "full length" polypeptide.

"Hit" refers to a sequence whose annotation will be used to describe a given template. Criteria for selecting the top hit are as follows: if the template has one or more exact nucleic acid matches, the top hit is the exact match with highest percent identity. If the template has no exact matches but has significant protein hits, the top hit is the protein hit with the lowest E-value. If the template has no significant protein hits, but does have significant non-exact nucleotide hits, the top hit is the nucleotide hit with the lowest E-value.

"Homology" refers to sequence similarity either between a reference nucleic acid sequence and at least a fragment of an mddt or between a reference amino acid sequence and a fragment of an MDDT.

"Hybridization" refers to the process by which a strand of nucleotides anneals with a complementary strand through base pairing. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under defined annealing conditions, and remain hybridized after the "washing" step. The defined hybridization conditions include the annealing conditions and the washing step(s), the latter of which is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid probes that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency.

Generally, stringency of hybridization is expressed with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization is well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, or 55°C may be used. SSC concentration may be varied from about 0.2 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 μg/ml. Useful variations on these conditions will be readily apparent to those skilled in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their resultant proteins.

Other parameters, such as temperature, salt concentration, and detergent concentration may be varied to achieve the desired stringency. Denaturants, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as RNA:DNA hybridizations. Appropriate hybridization conditions are routinely determinable by one of ordinary skill in the art.

"Immunogenic" describes the potential for a natural, recombinant, or synthetic peptide, epitope, polypeptide, or protein to induce antibody production in appropriate animals, cells, or cell lines.

"Insertion" or "addition" refers to a change in either a nucleic or amino acid sequence in which at least one nucleotide or residue, respectively, is added to the sequence.

"Labeling" refers to the covalent or noncovalent joining of a polynucleotide, polypeptide, or antibody with a reporter molecule capable of producing a detectable or measurable signal.

"Microarray" is any arrangement of nucleic acids, amino acids, antibodies, etc., on a substrate. The substrate may be a solid support such as beads, glass, paper, nitrocellulose, nylon, or an appropriate membrane.

"Linkers" are short stretches of nucleotide sequence which may be added to a vector or an mddt to create restriction endonuclease sites to facilitate cloning. "Polylinkers" are engineered to incorporate multiple restriction enzyme sites and to provide for the use of enzymes which leave 5' or 3' overhangs (e.g., BamHI, EcoRI, and HindIII) and those which provide blunt ends (e.g., EcoRV, SnaBI, and StuI).

"Naturally occurring" refers to an endogenous polynucleotide or polypeptide that may be isolated from viruses or prokaryotic or eukaryotic cells.

"Nucleic acid sequence" refers to the specific order of nucleotides joined by phosphodiester bonds in a linear, polymeric arrangement. Depending on the number of nucleotides, the nucleic acid sequence can be considered an oligomer, oligonucleotide, or polynucleotide. The nucleic acid can be DNA, RNA, or any nucleic acid analog, such as PNA, may be of genomic or synthetic origin, may be either double-stranded or single-stranded, and can represent either the sense or antisense (complementary) strand.

"Oligomer" refers to a nucleic acid sequence of at least about 6 nucleotides and as many as about 60 nucleotides, preferably about 15 to 40 nucleotides, and most preferably between about 20 and 30 nucleotides, that may be used in hybridization or amplification technologies. Oligomers may be used as, e.g., primers for PCR, and are usually chemically synthesized.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to a DNA mimic in which nucleotide bases are attached to a pseudopeptide backbone to increase stability. PNAs, also designated antigene agents, can prevent gene expression by targeting complementary messenger RNA.

The phrases "percent identity" and "% identity", as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and Sharp, P.M. (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to determine alignment between a known polynucleotide sequence and other sequences on a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2/. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in figures or Sequence Listings, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity", as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity of the substituted residue, thus preserving the structure (and therefore function) of the folded polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalty

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in figures or Sequence Listings, may be used to describe a length over which percentage identity may be measured.

"Post-translational modification" of an MDDT may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu and the MDDT.

"Probe" refers to mddt or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the figures and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000

nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

"Purified" refers to molecules, either polynucleotides or polypeptides that are isolated or separated from their natural environment and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other compounds with which they are naturally associated.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

"Regulatory element" refers to a nucleic acid sequence from nontranslated regions of a gene,

and includes enhancers, promoters, introns, and 3' untranslated regions, which interact with host proteins to carry out or regulate transcription or translation.

"Reporter" molecules are chemical or biochemical moieties used for labeling a nucleic acid, an amino acid, or an antibody. They include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

"Sample" is used in its broadest sense. Samples may contain nucleic or amino acids, antibodies, or other materials, and may be derived from any source (e.g., bodily fluids including, but not limited to, saliva, blood, and urine; chromosome(s), organelles, or membranes isolated from a cell; genomic DNA, RNA, or cDNA in solution or bound to a substrate; and cleared cells or tissues or blots or imprints from such cells or tissues).

"Specific binding" or "specifically binding" refers to the interaction between a protein or peptide and its agonist, antibody, antagonist, or other binding partner. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

"Substitution" refers to the replacement of at least one nucleotide or amino acid by a different nucleotide or amino acid.

"Substrate" refers to any suitable rigid or semi-rigid support including, e.g., membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles or capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular tissue or cell type under given conditions at a given time.

"Transformation" refers to a process by which exogenous DNA enters a recipient cell. Transformation may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the host cell being

transformed.

"Transformants" include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as cells which transiently express inserted DNA or RNA.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, and plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 25% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 30%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or even at least 98% or greater sequence identity over a certain defined length. The variant may result in "conservative" amino acid changes which do not affect structural and/or chemical properties. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

In an alternative, variants of the polynucleotides of the present invention may be generated through recombinant methods. One possible method is a DNA shuffling technique such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of MDDT, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

In a particular embodiment, cDNA sequences derived from human tissues and cell lines were aligned based on nucleotide sequence identity and assembled into "consensus" or "template" sequences which are designated by the template identification numbers (template IDs) in column 2 of Table 1. The sequence identification numbers (SEQ ID NO:s) corresponding to the template IDs are shown in column 1. The template sequences have similarity to GenBank sequences, or "hits," as designated by the GI Numbers in column 3. The statistical probability of each GenBank hit is indicated by a probability score in column 4, and the functional annotation corresponding to each GenBank hit is listed in column 5.

The invention incorporates the nucleic acid sequences of these templates as disclosed in the Sequence Listing and the use of these sequences in the diagnosis and treatment of disease states

characterized by defects in disease detection and treatment molecules. The invention further utilizes these sequences in hybridization and amplification technologies, and in particular, in technologies which assess gene expression patterns correlated with specific cells or tissues and their responses in vivo or in vitro to pharmaceutical agents, toxins, and other treatments. In this manner, the sequences of the present invention are used to develop a transcript image for a particular cell or tissue.

Derivation of Nucleic Acid Sequences

cDNA was isolated from libraries constructed using RNA derived from normal and diseased human tissues and cell lines. The human tissues and cell lines used for cDNA library construction were selected from a broad range of sources to provide a diverse population of cDNAs representative of gene transcription throughout the human body. Descriptions of the human tissues and cell lines used for cDNA library construction are provided in the LIFESEQ database (Incyte Genomics, Inc. (Incyte), Palo Alto CA). Human tissues were broadly selected from, for example, cardiovascular, dermatologic, endocrine, gastrointestinal, hematopoietic/immune system, musculoskeletal, neural, reproductive, and urologic sources.

Cell lines used for cDNA library construction were derived from, for example, leukemic cells, teratocarcinomas, neuroepitheliomas, cervical carcinoma, lung fibroblasts, and endothelial cells. Such cell lines include, for example, THP-1, Jurkat, HUVEC, hNT2, WI38, HeLa, and other cell lines commonly used and available from public depositories (American Type Culture Collection, Manassas VA). Prior to mRNA isolation, cell lines were untreated, treated with a pharmaceutical agent such as 5'-aza-2'-deoxycytidine, treated with an activating agent such as lipopolysaccharide in the case of leukocytic cell lines, or, in the case of endothelial cell lines, subjected to shear stress.

Sequencing of the cDNAs

Methods for DNA sequencing are well known in the art. Conventional enzymatic methods employ the Klenow fragment of DNA polymerase I, SEQUENASE DNA polymerase (U.S. Biochemical Corporation, Cleveland OH), Taq polymerase (Applied Biosystems, Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, Inc. (Amersham Pharmacia Biotech), Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies Inc. (Life Technologies), Gaithersburg MD), to extend the nucleic acid sequence from an oligonucleotide primer annealed to the DNA template of interest. Methods have been developed for the use of both single-stranded and double-stranded templates. Chain termination reaction products may be electrophoresed on urea-polyacrylamide gels and detected either by autoradiography (for radioisotope-labeled nucleotides) or by fluorescence (for

fluorophore-labeled nucleotides). Automated methods for mechanized reaction preparation, sequencing, and analysis using fluorescence detection methods have been developed. Machines used to prepare cDNAs for sequencing can include the MICROLAB 2200 liquid transfer system (Hamilton Company (Hamilton), Reno NV), Peltier thermal cycler (PTC200; MJ Research, Inc. (MJ Research), Watertown MA), and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing can be carried out using, for example, the ABI 373 or 377 (Applied Biosystems) or MEGABACE 1000 (Molecular Dynamics, Inc. (Molecular Dynamics), Sunnyvale CA) DNA sequencing systems, or other automated and manual sequencing systems well known in the art.

The nucleotide sequences of the Sequence Listing have been prepared by current, state-of-the-art, automated methods and, as such, may contain occasional sequencing errors or unidentified nucleotides. Such unidentified nucleotides are designated by an N. These infrequent unidentified bases do not represent a hindrance to practicing the invention for those skilled in the art. Several methods employing standard recombinant techniques may be used to correct errors and complete the missing sequence information. (See, e.g., those described in Ausubel, F.M. et al. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY; and Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY.)

Assembly of cDNA Sequences

Human polynucleotide sequences may be assembled using programs or algorithms well known in the art. Sequences to be assembled are related, wholly or in part, and may be derived from a single or many different transcripts. Assembly of the sequences can be performed using such programs as PHRAP (Phils Revised Assembly Program) and the GELVIEW fragment assembly system (GCG), or other methods known in the art.

Alternatively, cDNA sequences are used as "component" sequences that are assembled into "template" or "consensus" sequences as follows. Sequence chromatograms are processed, verified, and quality scores are obtained using PHRED. Raw sequences are edited using an editing pathway known as Block 1 (See, e.g., the LIFESEQ Assembled User Guide, Incyte Genomics, Palo Alto, CA). A series of BLAST comparisons is performed and low-information segments and repetitive elements (e.g., dinucleotide repeats, Alu repeats, etc.) are replaced by "n's", or masked, to prevent spurious matches. Mitochondrial and ribosomal RNA sequences are also removed. The processed sequences are then loaded into a relational database management system (RDMS) which assigns edited sequences to existing templates, if available. When additional sequences are added into the RDMS, a process is initiated which modifies existing templates or creates new templates from works in

progress (i.e., nonfinal assembled sequences) containing queued sequences or the sequences themselves. After the new sequences have been assigned to templates, the templates can be merged into bins. If multiple templates exist in one bin, the bin can be split and the templates reannotated.

Once gene bins have been generated based upon sequence alignments, bins are "clone joined" based upon clone information. Clone joining occurs when the 5' sequence of one clone is present in one bin and the 3' sequence from the same clone is present in a different bin, indicating that the two bins should be merged into a single bin. Only bins which share at least two different clones are merged.

A resultant template sequence may contain either a partial or a full length open reading frame, or all or part of a genetic regulatory element. This variation is due in part to the fact that the full length cDNAs of many genes are several hundred, and sometimes several thousand, bases in length. With current technology, cDNAs comprising the coding regions of large genes cannot be cloned because of vector limitations, incomplete reverse transcription of the mRNA, or incomplete "second strand" synthesis. Template sequences may be extended to include additional contiguous sequences derived from the parent RNA transcript using a variety of methods known to those of skill in the art. Extension may thus be used to achieve the full length coding sequence of a gene.

Analysis of the cDNA Sequences

The cDNA sequences are analyzed using a variety of programs and algorithms which are well known in the art. (See, e.g., Ausubel, 1997, supra, Chapter 7.7; Meyers, R.A. (Ed.) (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853; and Table 7.) These analyses comprise both reading frame determinations, e.g., based on triplet codon periodicity for particular organisms (Fickett, J.W. (1982) Nucleic Acids Res. 10:5303-5318); analyses of potential start and stop codons; and homology searches.

Computer programs known to those of skill in the art for performing computer-assisted searches for amino acid and nucleic acid sequence similarity, include, for example, Basic Local Alignment Search Tool (BLAST; Altschul, S.F. (1993) J. Mol. Evol. 36:290-300; Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410). BLAST is especially useful in determining exact matches and comparing two sequence fragments of arbitrary but equal lengths, whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user (Karlin, S. et al. (1988) Proc. Natl. Acad. Sci. USA 85:841-845). Using an appropriate search tool (e.g., BLAST or HMM), GenBank, SwissProt, BLOCKS, PFAM and other databases may be searched for sequences containing regions of homology to a query mddt or MDDT of the present invention.

Other approaches to the identification, assembly, storage, and display of nucleotide and polypeptide sequences are provided in "Relational Database for Storing Biomolecule Information,"

U.S.S.N. 08/947,845, filed October 9, 1997; "Project-Based Full-Length Biomolecular Sequence Database," U.S.S.N. 08/811,758, filed March 6, 1997; and "Relational Database and System for Storing Information Relating to Biomolecular Sequences," U.S.S.N. 09/034,807, filed March 4, 1998, all of which are incorporated by reference herein in their entirety.

Protein hierarchies can be assigned to the putative encoded polypeptide based on, e.g., motif, BLAST, or biological analysis. Methods for assigning these hierarchies are described, for example, in "Database System Employing Protein Function Hierarchies for Viewing Biomolecular Sequence Data," U.S.S.N. 08/812,290, filed March 6, 1997, incorporated herein by reference.

Human Disease Detection and Treatment Molecule Sequences

The mddt of the present invention may be used for a variety of diagnostic and therapeutic purposes. For example, an mddt may be used to diagnose a particular condition, disease, or disorder associated with disease detection and treatment molecules. Such conditions, diseases, and disorders include, but are not limited to, a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an autoimmune/inflammatory disorder, such as actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma. The mddt can be used to detect the presence of,

or to quantify the amount of, an mddt-related polynucleotide in a sample. This information is then compared to information obtained from appropriate reference samples, and a diagnosis is established. Alternatively, a polynucleotide complementary to a given mddt can inhibit or inactivate a therapeutically relevant gene related to the mddt.

Analysis of mddt Expression Patterns

The expression of mddt may be routinely assessed by hybridization-based methods to determine, for example, the tissue-specificity, disease-specificity, or developmental stage-specificity of mddt expression. For example, the level of expression of mddt may be compared among different cell types or tissues, among diseased and normal cell types or tissues, among cell types or tissues at different developmental stages, or among cell types or tissues undergoing various treatments. This type of analysis is useful, for example, to assess the relative levels of mddt expression in fully or partially differentiated cells or tissues, to determine if changes in mddt expression levels are correlated with the development or progression of specific disease states, and to assess the response of a cell or tissue to a specific therapy, for example, in pharmacological or toxicological studies. Methods for the analysis of mddt expression are based on hybridization and amplification technologies and include membrane-based procedures such as northern blot analysis, high-throughput procedures that utilize, for example, microarrays, and PCR-based procedures.

Hybridization and Genetic Analysis

The mddt, their fragments, or complementary sequences, may be used to identify the presence of and/or to determine the degree of similarity between two (or more) nucleic acid sequences. The mddt may be hybridized to naturally occurring or recombinant nucleic acid sequences under appropriately selected temperatures and salt concentrations. Hybridization with a probe based on the nucleic acid sequence of at least one of the mddt allows for the detection of nucleic acid sequences, including genomic sequences, which are identical or related to the mddt of the Sequence Listing. Probes may be selected from non-conserved or unique regions of at least one of the polynucleotides of SEQ ID NO:1-45 and tested for their ability to identify or amplify the target nucleic acid sequence using standard protocols.

Polynucleotide sequences that are capable of hybridizing, in particular, to those shown in SEQ ID NO:1-45 and fragments thereof, can be identified using various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions are discussed in "Definitions."

A probe for use in Southern or northern hybridization may be derived from a fragment of an

mddt sequence, or its complement, that is up to several hundred nucleotides in length and is either single-stranded or double-stranded. Such probes may be hybridized in solution to biological materials such as plasmids, bacterial, yeast, or human artificial chromosomes, cleared or sectioned tissues, or to artificial substrates containing mddt. Microarrays are particularly suitable for identifying the presence of and detecting the level of expression for multiple genes of interest by examining gene expression correlated with, e.g., various stages of development, treatment with a drug or compound, or disease progression. An array analogous to a dot or slot blot may be used to arrange and link polynucleotides to the surface of a substrate using one or more of the following: mechanical (vacuum), chemical, thermal, or UV bonding procedures. Such an array may contain any number of mddt and may be produced by hand or by using available devices, materials, and machines.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

Probes may be labeled by either PCR or enzymatic techniques using a variety of commercially available reporter molecules. For example, commercial kits are available for radioactive and chemiluminescent labeling (Amersham Pharmacia Biotech) and for alkaline phosphatase labeling (Life Technologies). Alternatively, mddt may be cloned into commercially available vectors for the production of RNA probes. Such probes may be transcribed in the presence of at least one labeled nucleotide (e.g., ^{32}P-ATP, Amersham Pharmacia Biotech).

Additionally the polynucleotides of SEQ ID NO:1-45 or suitable fragments thereof can be used to isolate full length cDNA sequences utilizing hybridization and/or amplification procedures well known in the art, e.g., cDNA library screening, PCR amplification, etc. The molecular cloning of such full length cDNA sequences may employ the method of cDNA library screening with probes using the hybridization, stringency, washing, and probing strategies described above and in Ausubel, supra, Chapters 3, 5, and 6. These procedures may also be employed with genomic libraries to isolate genomic sequences of mddt in order to analyze, e.g., regulatory elements.

Genetic Mapping

Gene identification and mapping are important in the investigation and treatment of almost all conditions, diseases, and disorders. Cancer, cardiovascular disease, Alzheimer's disease, arthritis, diabetes, and mental illnesses are of particular interest. Each of these conditions is more complex than the single gene defects of sickle cell anemia or cystic fibrosis, with select groups of genes being

predictive of predisposition for a particular condition, disease, or disorder. For example, cardiovascular disease may result from malfunctioning receptor molecules that fail to clear cholesterol from the bloodstream, and diabetes may result when a particular individual's immune system is activated by an infection and attacks the insulin-producing cells of the pancreas. In some studies, Alzheimer's disease has been linked to a gene on chromosome 21; other studies predict a different gene and location. Mapping of disease genes is a complex and reiterative process and generally proceeds from genetic linkage analysis to physical mapping.

As a condition is noted among members of a family, a genetic linkage map traces parts of chromosomes that are inherited in the same pattern as the condition. Statistics link the inheritance of particular conditions to particular regions of chromosomes, as defined by RFLP or other markers. (See, for example, Lander, E. S. and Botstein, D. (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.) Occasionally, genetic markers and their locations are known from previous studies. More often, however, the markers are simply stretches of DNA that differ among individuals. Examples of genetic linkage maps can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site.

In another embodiment of the invention, mddt sequences may be used to generate hybridization probes useful in chromosomal mapping of naturally occurring genomic sequences. Either coding or noncoding sequences of mddt may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of an mddt coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Meyers, supra, pp. 965-968.) Correlation between the location of mddt on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The mddt sequences may also be used to detect polymorphisms that are genetically linked to the inheritance of a particular condition, disease, or disorder.

In situ hybridization of chromosomal preparations and genetic mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending existing genetic

maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of the corresponding human chromosome is not known. These new marker sequences can be mapped to human chromosomes and may provide valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once a disease or syndrome has been crudely correlated by genetic linkage with a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequences of the subject invention may also be used to detect differences in chromosomal architecture due to translocation, inversion, etc., among normal, carrier, or affected individuals.

Once a disease-associated gene is mapped to a chromosomal region, the gene must be cloned in order to identify mutations or other alterations (e.g., translocations or inversions) that may be correlated with disease. This process requires a physical map of the chromosomal region containing the disease-gene of interest along with associated markers. A physical map is necessary for determining the nucleotide sequence of and order of marker genes on a particular chromosomal region. Physical mapping techniques are well known in the art and require the generation of overlapping sets of cloned DNA fragments from a particular organelle, chromosome, or genome. These clones are analyzed to reconstruct and catalog their order. Once the position of a marker is determined, the DNA from that region is obtained by consulting the catalog and selecting clones from that region. The gene of interest is located through positional cloning techniques using hybridization or similar methods.

Diagnostic Uses

The mddt of the present invention may be used to design probes useful in diagnostic assays. Such assays, well known to those skilled in the art, may be used to detect or confirm conditions, disorders, or diseases associated with abnormal levels of mddt expression. Labeled probes developed from mddt sequences are added to a sample under hybridizing conditions of desired stringency. In some instances, mddt, or fragments or oligonucleotides derived from mddt, may be used as primers in amplification steps prior to hybridization. The amount of hybridization complex formed is quantified and compared with standards for that cell or tissue. If mddt expression varies significantly from the standard, the assay indicates the presence of the condition, disorder, or disease. Qualitative or quantitative diagnostic methods may include northern, dot blot, or other membrane or dip-stick based technologies or multiple-sample format technologies such as PCR, enzyme-linked immunosorbent assay (ELISA)-like, pin, or chip-based assays.

The probes described above may also be used to monitor the progress of conditions, disorders, or diseases associated with abnormal levels of mddt expression, or to evaluate the efficacy of a particular therapeutic treatment. The candidate probe may be identified from the mddt that are specific to a given human tissue and have not been observed in GenBank or other genome databases. Such a probe may be used in animal studies, preclinical tests, clinical trials, or in monitoring the treatment of an individual patient. In a typical process, standard expression is established by methods well known in the art for use as a basis of comparison, samples from patients affected by the disorder or disease are combined with the probe to evaluate any deviation from the standard profile, and a therapeutic agent is administered and effects are monitored to generate a treatment profile. Efficacy is evaluated by determining whether the expression progresses toward or returns to the standard normal pattern. Treatment profiles may be generated over a period of several days or several months. Statistical methods well known to those skilled in the art may be use to determine the significance of such therapeutic agents.

The polynucleotides are also useful for identifying individuals from minute biological samples, for example, by matching the RFLP pattern of a sample's DNA to that of an individual's DNA. The polynucleotides of the present invention can also be used to determine the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, an individual can be identified through a unique set of DNA sequences. Once a unique ID database is established for an individual, positive identification of that individual can be made from extremely small tissue samples.

In a particular aspect, oligonucleotide primers derived from the mddt of the invention may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from mddt are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequences of individual overlapping DNA fragments which assemble into a common consensus

sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

DNA-based identification techniques are critical in forensic technology. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using, e.g., PCR, to identify individuals. (See, e.g., Erlich, H. (1992) PCR Technology, Freeman and Co., New York, NY). Similarly, polynucleotides of the present invention can be used as polymorphic markers.

There is also a need for reagents capable of identifying the source of a particular tissue. Appropriate reagents can comprise, for example, DNA probes or primers prepared from the sequences of the present invention that are specific for particular tissues. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

The polynucleotides of the present invention can also be used as molecular weight markers on nucleic acid gels or Southern blots, as diagnostic probes for the presence of a specific mRNA in a particular cell type, in the creation of subtracted cDNA libraries which aid in the discovery of novel polynucleotides, in selection and synthesis of oligomers for attachment to an array or other support, and as an antigen to elicit an immune response.

Disease Model Systems Using mddt

The mddt of the invention or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus

generated may be tested with potential therapeutic or toxic agents.

The mddt of the invention may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

The mddt of the invention can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of mddt is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress mddt, resulting, e.g., in the secretion of MDDT in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

Screening Assays

MDDT encoded by polynucleotides of the present invention may be used to screen for molecules that bind to or are bound by the encoded polypeptides. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the bound molecule. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a ligand or fragment thereof, a natural substrate, or a structural or functional mimetic. (See, Coligan et al., (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or to at least a fragment of the receptor, e.g., the active site. In either case, the molecule can be rationally designed using known techniques. Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing the polypeptide or cell membrane fractions which contain the expressed polypeptide are then contacted with a test compound and binding, stimulation, or inhibition of activity of either the polypeptide or the molecule is analyzed.

An assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. Alternatively, the assay may assess binding in the presence of a labeled competitor.

Additionally, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay using, e.g., a monoclonal or polyclonal antibody, can measure polypeptide level in a sample. The antibody can measure polypeptide level by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of the above assays can be used in a diagnostic or prognostic context. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Transcript Imaging and Toxicological Testing

Another embodiment relates to the use of mddt to develop a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity pertaining to disease detection and treatment molecules.

Transcript images which profile mddt expression may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect mddt expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile mddt expression may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are

indicative of mechanisms of action and toxicity (Nuwaysir, E. F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and Anderson, N. L. (2000) Toxicol. Lett. 112-113:467-71, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of MDDT encoded by polynucleotides of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is

generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for MDDT to quantify the levels of MDDT expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-11; Mendoze, L. G. et al. (1999) Biotechniques 27:778-88). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N. L. and Seilhamer, J. (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the MDDT encoded by polynucleotides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the MDDT encoded by polynucleotides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological

sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Transcript images may be used to profile mddt expression in distinct tissue types. This process can be used to determine disease detection and treatment molecule activity in a particular tissue type relative to this activity in a different tissue type. Transcript images may be used to generate a profile of mddt expression characteristic of diseased tissue. Transcript images of tissues before and after treatment may be used for diagnostic purposes, to monitor the progression of disease, and to monitor the efficacy of drug treatments for diseases which affect the activity of disease detection and treatment molecules.

Transcript images of cell lines can be used to assess disease detection and treatment molecule activity and/or to identify cell lines that lack or misregulate this activity. Such cell lines may then be treated with pharmaceutical agents, and a transcript image following treatment may indicate the efficacy of these agents in restoring desired levels of this activity. A similar approach may be used to assess the toxicity of pharmaceutical agents as reflected by undesirable changes in disease detection and treatment molecule activity. Candidate pharmaceutical agents may be evaluated by comparing their associated transcript images with those of pharmaceutical agents of known effectiveness.

Antisense Molecules

The polynucleotides of the present invention are useful in antisense technology. Antisense technology or therapy relies on the modulation of expression of a target protein through the specific binding of an antisense sequence to a target sequence encoding the target protein or directing its expression. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ; Alama, A. et al. (1997) Pharmacol. Res. 36(3):171-178; Crooke, S.T. (1997) Adv. Pharmacol. 40:1-49; Sharma, H.W. and R. Narayanan (1995) Bioessays 17(12):1055-1063; and Lavrosky, Y. et al. (1997) Biochem. Mol. Med. 62(1):11-22.) An antisense sequence is a polynucleotide sequence capable of specifically hybridizing to at least a portion of the target sequence. Antisense sequences bind to cellular mRNA and/or genomic DNA, affecting translation and/or transcription. Antisense sequences can be DNA, RNA, or nucleic acid mimics and analogs. (See, e.g., Rossi, J.J. et al. (1991) Antisense Res. Dev. 1(3):285-288; Lee, R. et al. (1998) Biochemistry 37(3):900-1010; Pardridge, W.M. et al. (1995) Proc. Natl. Acad. Sci. USA 92(12):5592-5596; and Nielsen, P. E. and Haaima, G. (1997) Chem. Soc. Rev. 96:73-78.) Typically, the binding which results in modulation of expression occurs through hybridization or binding of complementary base pairs. Antisense sequences can also bind to DNA duplexes through specific interactions in the major groove of the double helix.

The polynucleotides of the present invention and fragments thereof can be used as antisense sequences to modify the expression of the polypeptide encoded by mddt. The antisense sequences can be produced ex vivo, such as by using any of the ABI nucleic acid synthesizer series (Applied Biosystems) or other automated systems known in the art. Antisense sequences can also be produced biologically, such as by transforming an appropriate host cell with an expression vector containing the sequence of interest. (See, e.g., Agrawal, supra.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E., et al. (1998) J. Allergy Clin. Immunol. 102(3):469-475; and Scanlon, K.J., et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

Expression

In order to express a biologically active MDDT, the nucleotide sequences encoding MDDT or fragments thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding MDDT and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, supra, Chapters 4, 8, 16, and 17; and Ausubel, supra, Chapters 9, 10, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding MDDT. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or

animal (mammalian) cell systems. (See, e.g., Sambrook, supra; Ausubel, 1995, supra, Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al., (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

For long term production of recombinant proteins in mammalian systems, stable expression of MDDT in cell lines is preferred. For example, sequences encoding MDDT can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Any number of selection systems may be used to recover transformed cell lines. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.; Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14; Hartman, S.C. and R.C.Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051; Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Therapeutic Uses of mddt

The mddt of the invention may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor

VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and Somia, N. (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in mddt expression or regulation causes disease, the expression of mddt from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in mddt are treated by constructing mammalian expression vectors comprising mddt and introducing these vectors by mechanical means into mddt-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and Anderson, W.F. (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and Récipon, H. (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of mddt include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). The mddt of the invention may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and Bujard, H. (1992) Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551; Gossen, M. et al., (1995) Science 268:1766-1769; Rossi, F.M.V. and Blau, H.M. (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. supra), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding MDDT from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental

parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and Eb, A.J. (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to mddt expression are treated by constructing a retrovirus vector consisting of (i) mddt under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. U.S.A. 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and Miller, A.D. (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. U.S.A. 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver mddt to cells which have one or more genetic abnormalities with respect to the expression of mddt. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and Somia, N. (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver mddt to target cells which have one or more genetic abnormalities with respect to the expression of mddt. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing mddt to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res.169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W. F. et al. 1999 J. Virol. 73:519-532 and Xu, H. et al., (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver mddt to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and Li, K-J. (1998) Curr. Opin. Biotech. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting mddt into the alphavirus genome in place of the capsid-coding region results in the production of a large number of mddt RNAs and the synthesis of high levels of MDDT in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of mddt into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA

transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Antibodies

Anti-MDDT antibodies may be used to analyze protein expression levels. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, and Fab fragments. For descriptions of and protocols of antibody technologies, see, e.g., Pound J.D. (1998) Immunochemical Protocols, Humana Press, Totowa, NJ.

The amino acid sequence encoded by the mddt of the Sequence Listing may be analyzed by appropriate software (e.g., LASERGENE NAVIGATOR software, DNASTAR) to determine regions of high immunogenicity. The optimal sequences for immunization are selected from the C-terminus, the N-terminus, and those intervening, hydrophilic regions of the polypeptide which are likely to be exposed to the external environment when the polypeptide is in its natural conformation. Analysis used to select appropriate epitopes is also described by Ausubel (1997, supra, Chapter 11.7). Peptides used for antibody induction do not need to have biological activity; however, they must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least five amino acids, preferably at least 10 amino acids, and most preferably at least 15 amino acids. A peptide which mimics an antigenic fragment of the natural polypeptide may be fused with another protein such as keyhole hemolimpet cyanin (KLH; Sigma, St. Louis MO) for antibody production. A peptide encompassing an antigenic region may be expressed from an mddt, synthesized as described above, or purified from human cells.

Procedures well known in the art may be used for the production of antibodies. Various hosts including mice, goats, and rabbits, may be immunized by injection with a peptide. Depending on the host species, various adjuvants may be used to increase immunological response.

In one procedure, peptides about 15 residues in length may be synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using fmoc-chemistry and coupled to KLH (Sigma) by reaction with M-maleimidobenzoyl-N-hydroxysuccinimide ester (Ausubel, 1995, supra). Rabbits are immunized with the peptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity by binding the peptide to plastic, blocking with 1% bovine serum albumin (BSA), reacting with rabbit antisera, washing, and reacting with radioiodinated goat anti-rabbit IgG. Antisera with antipeptide activity are tested for anti-MDDT activity using protocols well known in the art, including ELISA, radioimmunoassay (RIA), and immunoblotting.

In another procedure, isolated and purified peptide may be used to immunize mice (about 100 µg of peptide) or rabbits (about 1 mg of peptide). Subsequently, the peptide is radioiodinated and used

to screen the immunized animals' B-lymphocytes for production of antipeptide antibodies. Positive cells are then used to produce hybridomas using standard techniques. About 20 mg of peptide is sufficient for labeling and screening several thousand clones. Hybridomas of interest are detected by screening with radioiodinated peptide to identify those fusions producing peptide-specific monoclonal antibody. In a typical protocol, wells of a multi-well plate (FAST, Becton-Dickinson, Palo Alto, CA) are coated with affinity-purified, specific rabbit-anti-mouse (or suitable anti-species IgG) antibodies at 10 mg/ml. The coated wells are blocked with 1% BSA and washed and exposed to supernatants from hybridomas. After incubation, the wells are exposed to radiolabeled peptide at 1 mg/ml.

Clones producing antibodies bind a quantity of labeled peptide that is detectable above background. Such clones are expanded and subjected to 2 cycles of cloning. Cloned hybridomas are injected into pristane-treated mice to produce ascites, and monoclonal antibody is purified from the ascitic fluid by affinity chromatography on protein A (Amersham Pharmacia Biotech). Several procedures for the production of monoclonal antibodies, including in vitro production, are described in Pound (supra). Monoclonal antibodies with antipeptide activity are tested for anti-MDDT activity using protocols well known in the art, including ELISA, RIA, and immunoblotting.

Antibody fragments containing specific binding sites for an epitope may also be generated. For example, such fragments include, but are not limited to, the F(ab')2 fragments produced by pepsin digestion of the antibody molecule, and the Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, construction of Fab expression libraries in filamentous bacteriophage allows rapid and easy identification of monoclonal fragments with desired specificity (Pound, supra, Chaps. 45-47). Antibodies generated against polypeptide encoded by mddt can be used to purify and characterize full-length MDDT protein and its activity, binding partners, etc.

Assays Using Antibodies

Anti-MDDT antibodies may be used in assays to quantify the amount of MDDT found in a particular human cell. Such assays include methods utilizing the antibody and a label to detect expression level under normal or disease conditions. The peptides and antibodies of the invention may be used with or without modification or labeled by joining them, either covalently or noncovalently, with a reporter molecule.

Protocols for detecting and measuring protein expression using either polyclonal or monoclonal antibodies are well known in the art. Examples include ELISA, RIA, and fluorescent activated cell sorting (FACS). Such immunoassays typically involve the formation of complexes between the MDDT and its specific antibody and the measurement of such complexes. These and other assays are described in Pound (supra).

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/185,213, U.S. Ser. No. 60/205,285, U.S. Ser. No. 60/205,232, U.S. Ser. No. 60/205,323, U.S. Ser. No. 60/205,287, U.S. Ser. No. 60/205,324, and U.S. Ser. No. 60/205,286, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from CLONTECH Laboratories, Inc. (Palo Alto CA) or isolated from various tissues. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In most cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega Corporation (Promega), Madison WI), OLIGOTEX latex particles (QIAGEN, Inc. (QIAGEN), Valencia CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Inc., Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene Cloning Systems, Inc. (Stratagene), La Jolla CA) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, Chapters 5.1 through 6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid

(Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: the Magic or WIZARD Minipreps DNA purification system (Promega); the AGTC Miniprep purification kit (Edge BioSystems, Gaithersburg MD); and the QIAWELL 8, QIAWELL 8 Plus, and QIAWELL 8 Ultra plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit (QIAGEN). Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format. (Rao, V.B. (1994) Anal. Biochem. 216:1-14.) Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Inc. (Molecular Probes), Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 thermal cycler (Applied Biosystems) or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific Corp., Sunnyvale CA) or the MICROLAB 2200 liquid transfer system (Hamilton). cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, Chapter 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

IV. Assembly and Analysis of Sequences

Component sequences from chromatograms were subject to PHRED analysis and assigned a quality score. The sequences having at least a required quality score were subject to various pre-processing editing pathways to eliminate, e.g., low quality 3' ends, vector and linker sequences, polyA tails, Alu repeats, mitochondrial and ribosomal sequences, bacterial contamination sequences, and sequences smaller than 50 base pairs. In particular, low-information sequences and repetitive elements (e.g., dinucleotide repeats, Alu repeats, etc.) were replaced by "n's", or masked, to prevent spurious matches.

Processed sequences were then subject to assembly procedures in which the sequences were assigned to gene bins (bins). Each sequence could only belong to one bin. Sequences in each gene bin were assembled to produce consensus sequences (templates). Subsequent new sequences were added to existing bins using BLASTn (v.1.4 WashU) and CROSSMATCH. Candidate pairs were identified as all BLAST hits having a quality score greater than or equal to 150. Alignments of at least 82% local identity were accepted into the bin. The component sequences from each bin were assembled using a version of PHRAP. Bins with several overlapping component sequences were assembled using DEEP PHRAP. The orientation (sense or antisense) of each assembled template was determined based on the number and orientation of its component sequences. Template sequences as disclosed in the sequence listing correspond to sense strand sequences (the "forward" reading frames), to the best determination. The complementary (antisense) strands are inherently disclosed herein. The component sequences which were used to assemble each template consensus sequence are listed in Table 4, along with their positions along the template nucleotide sequences.

Bins were compared against each other and those having local similarity of at least 82% were combined and reassembled. Reassembled bins having templates of insufficient overlap (less than 95% local identity) were re-split. Assembled templates were also subject to analysis by STITCHER/EXON MAPPER algorithms which analyze the probabilities of the presence of splice variants, alternatively spliced exons, splice junctions, differential expression of alternative spliced genes across tissue types or disease states, etc. These resulting bins were subject to several rounds of the above assembly procedures.

Once gene bins were generated based upon sequence alignments, bins were clone joined based upon clone information. If the 5' sequence of one clone was present in one bin and the 3' sequence from the same clone was present in a different bin, it was likely that the two bins actually belonged together in a single bin. The resulting combined bins underwent assembly procedures to regenerate the consensus sequences.

The final assembled templates were subsequently annotated using the following procedure. Template sequences were analyzed using BLASTn (v2.0, NCBI) versus gbpri (GenBank version 120). "Hits" were defined as an exact match having from 95% local identity over 200 base pairs through 100% local identity over 100 base pairs, or a homolog match having an E-value, i.e. a probability score, of $\leq 1 \times 10^{-8}$. The hits were subject to frameshift FASTx versus GENPEPT (GenBank version 120). (See Table 7). In this analysis, a homolog match was defined as having an E-value of $\leq 1 \times 10^{-8}$. The assembly method used above was described in "System and Methods for Analyzing Biomolecular Sequences," U.S.S.N. 09/276,534, filed March 25, 1999, and the LIFESEQ Gold user manual (Incyte) both incorporated by reference herein.

Following assembly, template sequences were subjected to motif, BLAST, and functional analyses, and categorized in protein hierarchies using methods described in, e.g., "Database System Employing Protein Function Hierarchies for Viewing Biomolecular Sequence Data," U.S.S.N. 08/812,290, filed March 6, 1997; "Relational Database for Storing Biomolecule Information," U.S.S.N. 08/947,845, filed October 9, 1997; "Project-Based Full-Length Biomolecular Sequence Database," U.S.S.N. 08/811,758, filed March 6, 1997; and "Relational Database and System for Storing Information Relating to Biomolecular Sequences," U.S.S.N. 09/034,807, filed March 4, 1998, all of which are incorporated by reference herein.

The template sequences were further analyzed by translating each template in all three forward reading frames and searching each translation against the Pfam database of hidden Markov model-based protein families and domains using the HMMER software package (available to the public from Washington University School of Medicine, St. Louis MO). Regions of templates which, when translated, contain similarity to Pfam consensus sequences are reported in Table 2, along with descriptions of Pfam protein domains and families. Only those Pfam hits with an E-value of $\leq 1 \times 10^{-3}$ are reported. (See also World Wide Web site http://pfam.wustl.edu/ for detailed descriptions of Pfam protein domains and families.)

Additionally, the template sequences were translated in all three forward reading frames, and each translation was searched against hidden Markov models for signal peptides using the HMMER software package. Construction of hidden Markov models and their usage in sequence analysis has been described. (See, for example, Eddy, S.R. (1996) Curr. Opin. Str. Biol. 6:361-365.) Only those signal peptide hits with a cutoff score of 11 bits or greater are reported. A cutoff score of 11 bits or greater corresponds to at least about 91-94% true-positives in signal peptide prediction. Template sequences were also translated in all three forward reading frames, and each translation was searched against TMAP, a program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation, with respect to the cell cytosol (Persson, B. and P. Argos (1994) J.

Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.) Regions of templates which, when translated, contain similarity to signal peptide or transmembrane consensus sequences are reported in Table 3.

The results of HMMER analysis as reported in Tables 2 and 3 may support the results of BLAST analysis as reported in Table 1 or may suggest alternative or additional properties of template-encoded polypeptides not previously uncovered by BLAST or other analyses.

Template sequences are further analyzed using the bioinformatics tools listed in Table 7, or using sequence analysis software known in the art such as MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Template sequences may be further queried against public databases such as the GenBank rodent, mammalian, vertebrate, prokaryote, and eukaryote databases.

The template sequences were translated to derive the corresponding longest open reading frame as presented by the polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues within the full length translated polypeptide. Polypeptide sequences were subsequently analyzed by querying against the GenBank protein database (GENPEPT, (GenBank version 121)). Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 6 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (GENPEPT) database. Column 1 shows the polypeptide sequence identification number (SEQ ID NO:) for the polypeptide segments of the invention. Column 2 shows the reading frame used in the translation of the polynucleotide sequences encoding the polypeptide segments. Column 3 shows the length of the translated polypeptide segments. Columns 4 and 5 show the start and stop nucleotide positions of the polynucleotide sequences encoding the polypeptide segments. Column 6 shows the GenBank identification number (GI Number) of the nearest GenBank homolog. Column 7 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 8 shows the annotation of the GenBank homolog.

V. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score x Percent Identity}}{5 \times \text{minimum } \{\text{length(Seq. 1), length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

VI. Tissue Distribution Profiling

A tissue distribution profile is determined for each template by compiling the cDNA library tissue classifications of its component cDNA sequences. Each component sequence, is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. Template sequences, component sequences, and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

Table 5 shows the tissue distribution profile for the templates of the invention. For each template, the three most frequently observed tissue categories are shown in column 3, along with the

percentage of component sequences belonging to each category. Only tissue categories with percentage values of ≥10% are shown. A tissue distribution of "widely distributed" in column 3 indicates percentage values of <10% in all tissue categories.

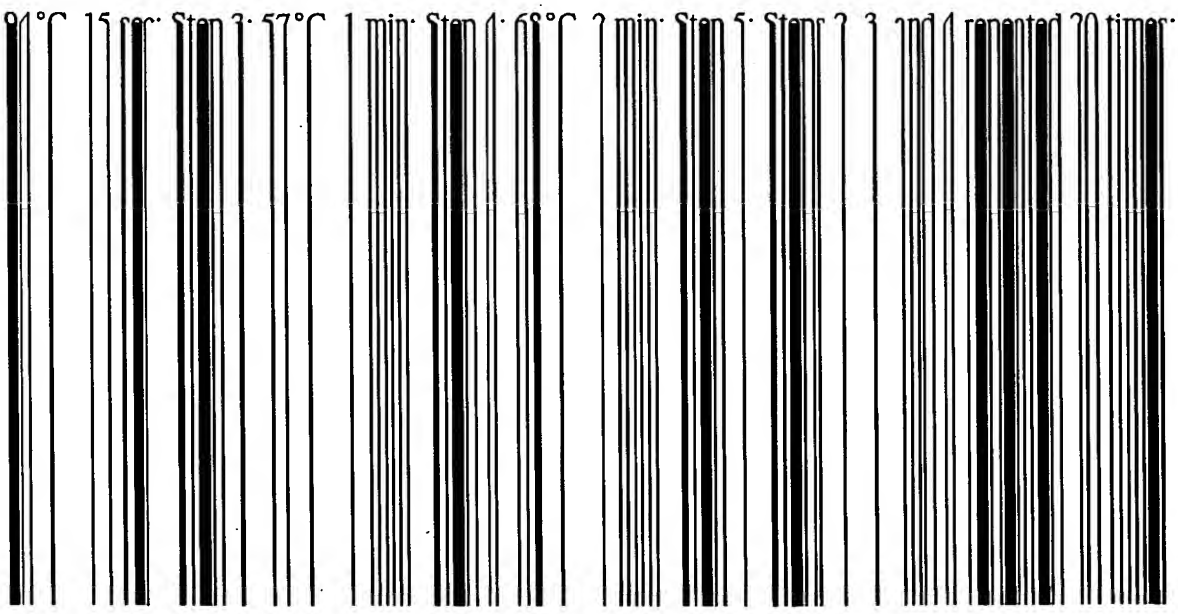
VII. Transcript Image Analysis

Transcript images are generated as described in Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, incorporated herein by reference.

VIII. Extension of Polynucleotide Sequences and Isolation of a Full-length cDNA

Oligonucleotide primers designed using an mddt of the Sequence Listing are used to extend the nucleic acid sequence. One primer is synthesized to initiate 5' extension of the template, and the other primer, to initiate 3' extension of the template. The initial primers may be designed using OLIGO 4.06 software (National Biosciences, Inc. (National Biosciences), Plymouth MN), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations are avoided. Selected human cDNA libraries are used to extend the sequence. If more than one extension is necessary or desired, additional or nested sets of primers are designed.

High fidelity amplification is obtained by PCR using methods well known in the art. PCR is performed in 96-well plates using the PTC-200 thermal cycler (MJ Research). The reaction mix contains DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and ß-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ are as follows: Step 1: 94°C, 3 min; Step 2:



to determine which reactions are successful in extending the sequence.

The extended nucleotides are desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides are separated on low concentration (0.6 to 0.8%) agarose gels, fragments are excised, and agar digested with AGAR ACE (Promega). Extended clones are religated using T4 ligase (New England Biolabs, Inc., Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells are selected on antibiotic-containing media, individual colonies are picked and cultured overnight at 37°C in 384-well plates in LB/2x carbenicillin liquid media.

The cells are lysed, and DNA is amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA is quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries are reamplified using the same conditions as described above. Samples are diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, the mddt is used to obtain regulatory sequences (promoters, introns, and enhancers) using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Labeling of Probes and Southern Hybridization Analyses

Hybridization probes derived from the mddt of the Sequence Listing are employed for screening cDNAs, mRNAs, or genomic DNA. The labeling of probe nucleotides between 100 and 1000 nucleotides in length is specifically described, but essentially the same procedure may be used with larger cDNA fragments. Probe sequences are labeled at room temperature for 30 minutes using a T4 polynucleotide kinase, γ^{32}P-ATP, and 0.5X One-Phor-All Plus (Amersham Pharmacia Biotech) buffer and purified using a ProbeQuant G-50 Microcolumn (Amersham Pharmacia Biotech). The probe mixture is diluted to 10^7 dpm/μg/ml hybridization buffer and used in a typical membrane-based hybridization analysis.

The DNA is digested with a restriction endonuclease such as Eco RV and is electrophoresed

through a 0.7% agarose gel. The DNA fragments are transferred from the agarose to nylon membrane (NYTRAN Plus, Schleicher & Schuell, Inc., Keene NH) using procedures specified by the manufacturer of the membrane. Prehybridization is carried out for three or more hours at 68°C, and hybridization is carried out overnight at 68°C. To remove non-specific signals, blots are sequentially washed at room temperature under increasingly stringent conditions, up to 0.1x saline sodium citrate (SSC) and 0.5% sodium dodecyl sulfate. After the blots are placed in a PHOSPHORIMAGER cassette (Molecular Dynamics) or are exposed to autoradiography film, hybridization patterns of standard and experimental lanes are compared. Essentially the same procedure is employed when screening RNA.

X. Chromosome Mapping of mddt

The cDNA sequences which were used to assemble SEQ ID NO:1-45 are compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that match SEQ ID NO:1-45 are assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as PHRAP (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon are used to determine if any of the clustered sequences have been previously mapped. Inclusion of a mapped sequence in a cluster will result in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location. The genetic map locations of SEQ ID NO:1-45 are described as ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters.

XI. Microarray Analysis

Probe Preparation from Tissue or Cell Samples

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and polyA$^+$ RNA is purified using the oligo (dT) cellulose method. Each polyA$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/µl oligo-dT primer (21mer), 1X first strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription

reaction is performed in a 25 ml volume containing 200 ng polyA$^+$ RNA with GEMBRIGHT kits (Incyte). Specific control polyA$^+$ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA (W. Lei, unpublished). As quantitative controls, the control mRNAs at 0.002 ng, 0.02 ng, 0.2 ng, and 2 ng are diluted into reverse transcription reaction at ratios of 1:100,000, 1:10,000, 1:1000, 1:100 (w/w) to sample mRNA respectively. The control mRNAs are diluted into reverse transcription reaction at ratios of 1:3, 3:1, 1:10, 10:1, 1:25, 25:1 (w/w) to sample mRNA differential expression patterns. After incubation at 37° C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85° C to the stop the reaction and degrade the RNA. Probes are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The probe is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester, PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford, MA) for 30 minutes at 60° C followed by washes in 0.2%

SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 µl of probe mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The probe mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 µl of 5x SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60° C. The arrays are washed for 10 min at 45° C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45° C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the probe mix at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two probes from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the

calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood, MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XII. Complementary Nucleic Acids

Sequences complementary to the mddt are used to detect, decrease, or inhibit expression of the naturally occurring nucleotide. The use of oligonucleotides comprising from about 15 to 30 base pairs is typical in the art. However, smaller or larger sequence fragments can also be used. Appropriate oligonucleotides are designed from the mddt using OLIGO 4.06 software (National Biosciences) or other appropriate programs and are synthesized using methods standard in the art or ordered from a commercial supplier. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent transcription factor binding to the promoter sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding and processing of the transcript.

XIII. Expression of MDDT

Expression and purification of MDDT is accomplished using bacterial or virus-based expression systems. For expression of MDDT in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express MDDT upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of MDDT in eukaryotic cells is achieved by infecting insect

or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding MDDT by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See e.g., Engelhard, supra; and Sandig, supra.)

In most expression systems, MDDT is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from MDDT at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak Company, Rochester NY). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, Chapters 10 and 16). Purified MDDT obtained by these methods can be used directly in the following activity assay.

XIV. Demonstration of MDDT Activity

MDDT, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled MDDT, washed, and any wells with labeled MDDT complex are assayed. Data obtained using different concentrations of MDDT are used to calculate values for the number, affinity, and association of MDDT with the candidate molecules.

Alternatively, molecules interacting with MDDT are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) Nature 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (CLONTECH).

MDDT may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions

between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XV. Functional Assays

MDDT function is assessed by expressing mddt at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen Corporation, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected.

Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; CLONTECH), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties.

FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of MDDT on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding MDDT and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Inc., Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art.

Expression of mRNA encoding MDDT and other genes of interest can be analyzed by northern analysis or microarray techniques.

XVI. Production of Antibodies

MDDT substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the MDDT amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding peptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, Chapter 11.)

Typically, peptides 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using fmoc-chemistry and coupled to KLH (Sigma) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, supra.) Rabbits are immunized with the peptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG. Antisera with antipeptide activity are tested for anti-MDDT activity using protocols well known in the art, including ELISA, RIA, and immunoblotting.

XVII. Purification of Naturally Occurring MDDT Using Specific Antibodies

Naturally occurring or recombinant MDDT is substantially purified by immunoaffinity chromatography using antibodies specific for MDDT. An immunoaffinity column is constructed by covalently coupling anti-MDDT antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing MDDT are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of MDDT (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/MDDT binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and MDDT is collected.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.

TABLE 1

SEQ ID NO:	Template ID	GI Number	Probability Score	Annotation
1	LG:977683.1:2000FEB18	g10764778	0	phosphoinositol 3-phosphate-binding protein-2 (Homo
2	LG:893050.1:2000FEB18	g6634025	2.00E-81	KIAA0379 protein (Homo sapiens)
3	LG:980153.1:2000FEB18	g7263990	0	dJ93K22.1 (novel protein (contains DKFZP564B116)) (Homo sapiens)
4	LG:350398.1:2000FEB18	g3882175	3.00E-10	KIAA0727 protein (Homo sapiens)
5	LG:475551.1:2000FEB18	g861029	0	SH3 domain binding protein (Mus musculus)
6	LG:481407.2:2000FEB18	g6119546	1.00E-41	hypothetical protein; 114721-113936 (Arabidopsis thaliana)
7	LI:443580.1:2000FEB01	g4589566	3.00E-34	KIAA0961 protein (Homo sapiens)
8	LI:803015.1:2000FEB01	g5262560	2.00E-35	hypothetical protein (Homo sapiens)
9	LG:027410.3:2000MAY19	g10438267	1.00E-65	unnamed protein product (Homo sapiens)
10	LG:171377.1:2000MAY19	g3077703	1.00E-107	mitsugumin29 (Oryctolagus cuniculus)
11	LG:352559.1:2000MAY19	g7243243	2.00E-43	KIAA1431 protein (Homo sapiens)
12	LG:247384.1:2000MAY19	g9945010	1.00E-118	RING-finger protein MURF (Mus musculus)
13	LG:403872.1:2000MAY19	g7020303	0	unnamed protein product (Homo sapiens)
14	LG:1135213.1:2000MAY19	g6692607	2.00E-65	MGA protein (Mus musculus)
15	LG:474284.2:2000MAY19	g1488047	2.00E-30	RING finger protein (Xenopus laevis)
16	LG:342147.1:2000MAY19	g2477511	3.00E-41	Homo sapiens p20 protein (pir B53814)
17	LG:1097300.1:2000MAY19	g2078531	1.00E-70	Mlark (Mus musculus)
18	LG:444850.9:2000MAY19	g199000	0	interferon-gamma inducible protein (Mus musculus)
19	LG:402231.6:2000MAY19	g7020737	6.00E-77	unnamed protein product (Homo sapiens)
20	LG:1076157.1:2000MAY19	g5262560	3.00E-65	hypothetical protein (Homo sapiens)
21	LG:1083142.1:2000MAY19	g4589566	3.00E-23	KIAA0961 protein (Homo sapiens)
22	LG:1083264.1:2000MAY19	g10047297	2.00E-25	KIAA1611 protein (Homo sapiens)
23	LG:350793.2:2000MAY19	g7242973	0	KIAA1309 protein (Homo sapiens)
24	LG:408751.3:2000MAY19	g8886025	1.00E-134	collapsin response mediator protein-5 (Homo sapiens)
25	LI:336120.1:2000MAY01	g1864085	1.00E-160	glypican-5 (Homo sapiens)
26	LI:234104.2:2000MAY01	g1518505	1.00E-114	G-protein coupled inwardly rectifying K+ channel (Mus musculus)
27	LI:450887.1:2000MAY01	g7629994	3.00E-34	60S RIBOSOMAL PROTEIN L36 homolog (Arabidopsis thaliana)
28	LI:119992.3:2000MAY01	g7243089	0	KIAA1354 protein (Homo sapiens)
29	LI:197241.2:2000MAY01	g7263990	0	dJ93K22.1 (novel protein (contains DKFZP564B116)) (Homo sapiens)
30	LI:406860.20:2000MAY01	g10435919	3.00E-57	unnamed protein product (Homo sapiens)

31	LI:142384.1:2000MAY01	g10436290	1.00E-131	unnamed protein product (Homo sapiens)
32	LI:895427.1:2000MAY01	g3184264	1.00E-106	F02569_2 (Homo sapiens)
33	LI:757439.1:2000MAY01	g7670362	1.00E-116	unnamed protein product (Mus musculus)
34	LI:1144066.1:2000MAY01	g3882281	7.00E-79	KIAA0780 protein (Homo sapiens)
35	LI:243660.4:2000MAY01	g4210501	0	BC85722_1 (Homo sapiens)
36	LI:334386.1:2000MAY01	g6330617	0	KIAA1223 protein (Homo sapiens)
37	LI:347572.1:2000MAY01	g9802433	1.00E-101	ACE-related carboxypeptidase ACE2 (Homo sapiens)
38	LI:817314.1:2000MAY01	g5802615	0	transient receptor potential 4 (Homo sapiens)
39	LI:000290.1:2000MAY01	g7242977	2.00E-51	KIAA1311 protein (Homo sapiens)
40	LI:023518.3:2000MAY01	g736727	2.00E-74	32 kd accessory protein (Bos taurus)
41	LI:1084246.1:2000MAY01	g5457031	0	protocadherin beta 12 (Homo sapiens)
42	LI:1165828.1:2000MAY01	g5457019	0	protocadherin alpha 7 short form protein (Homo sapiens)
43	LI:007302.1:2000MAY01	g5006250	0	TLR6 (Mus musculus)
44	LI:236386.4:2000MAY01	g6164628	1.00E-63	SH3 and PX domain-containing protein SH3PX1 (Homo sapiens)
45	LI:252904.5:2000MAY01	g7022971	2.00E-62	unnamed protein product (Homo sapiens)

TABLE 2

SEQ ID NO:	Template ID	Start	Stop	Frame	Pfam Hit	Pfam Description	E-value
1	LG:977683.1:2000FEB18	540	695	forward 3	PH	PH domain	6.70E-11
1	LG:977683.1:2000FEB18	204	293	forward 3	WW	WW domain	7.50E-05
2	LG:893050.1:2000FEB18	211	309	forward 1	ank	Ank repeat	1.60E-05
3	LG:980153.1:2000FEB18	754	852	forward 1	ank	Ank repeat	8.00E-04
3	LG:980153.1:2000FEB18	2131	2565	forward 1	BTB	BTB/POZ domain	6.90E-07
3	LG:980153.1:2000FEB18	1084	1239	forward 1	RCC1	Regulator of chromosome condensation	3.70E-04
4	LG:350398.1:2000FEB18	7	123	forward 1	myosin_head	Myosin head (motor domain)	2.60E-16
5	LG:475551.1:2000FEB18	702	1157	forward 3	RhoGAP	RhoGAP domain	8.10E-71
6	LG:481407.2:2000FEB18	225	440	forward 3	rrm	RNA recognition motif. (a.k.a. RRM, RBD	1.50E-22
6	LG:481407.2:2000FEB18	504	557	forward 3	zf-CCHC	Zinc knuckle	7.00E-04
7	LI:443580.1:2000FEB01	262	450	forward 1	KRAB	KRAB box	1.60E-41
7	LI:443580.1:2000FEB01	625	693	forward 1	zf-C2H2	Zinc finger, C2H2 type	2.20E-06
8	LI:803015.1:2000FEB01	159	299	forward 3	KRAB	KRAB box	2.30E-17
9	LG:027410.3:2000MAY19	177	290	forward 3	WD40	WD domain, G-beta repeat	6.20E-06
10	LG:171377.1:2000MAY19	300	848	forward 3	Synaptophysin	Synaptophysin / synaptoporin	2.10E-20
11	LG:352559.1:2000MAY19	125	313	forward 2	KRAB	KRAB box	1.60E-41
12	LG:247384.1:2000MAY19	182	256	forward 2	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	1.80E-06
13	LG:403872.1:2000MAY19	717	1187	forward 3	PAP2	PAP2 superfamily	1.80E-09
14	LG:1135213.1:2000MAY19	340	531	forward 1	T-box	T-box	8.80E-27
15	LG:474284.2:2000MAY19	73	195	forward 1	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	1.20E-13
16	LG:342147.1:2000MAY19	290	469	forward 2	crystallin	Alpha crystallin A chain, N terminal	3.10E-09
16	LG:342147.1:2000MAY19	452	628	forward 2	HSP20	Hsp20/alpha crystallin family	7.20E-12
17	LG:1097300.1:2000MAY19	59	250	forward 2	rrm	RNA recognition motif. (a.k.a. RRM, RBD	4.10E-16
18	LG:444850.9:2000MAY19	190	1290	forward 1	GBP	Guanylate-binding protein	4.20E-247
19	LG:402231.6:2000MAY19	258	380	forward 3	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	4.30E-05
20	LG:1076157.1:2000MAY19	180	320	forward 3	KRAB	KRAB box	3.40E-18
21	LG:1083142.1:2000MAY19	129	320	forward 3	KRAB	KRAB box	2.00E-42
22	LG:1083264.1:2000MAY19	440	628	forward 2	KRAB	KRAB box	2.30E-33
23	LG:350793.2:2000MAY19	570	722	forward 3	Kelch	Kelch motif	2.70E-11
24	LG:408751.3:2000MAY19	194	1051	forward 2	Dihydrooratase	Dihydroorotase-like	5.50E-07
25	LI:336120.1:2000MAY01	232	1398	forward 1	Glypican	Glypican	9.90E-141
25	LI:336120.1:2000MAY01	1476	1907	forward 3	Glypican	Glypican	8.60E-70
25	LI:336120.1:2000MAY01	503	775	forward 2	Glypican	Glypican	3.50E-46
26	LI:234104.2:2000MAY01	2517	3002	forward 3	IRK	Inward rectifier potassium channel	8.70E-111

26	LI:234104.2:2000MAY01	2965	3507	forward 1	IRK	Inward rectifier potassium channel	9.20E-111
27	LI:450887.1:2000MAY01	48	344	forward 3	Ribosomal_L36e	Ribosomal protein L36e	6.90E-41
28	LI:119992.3:2000MAY01	788	925	forward 2	Kelch	Kelch motif	1.50E-09
29	LI:197241.2:2000MAY01	1243	1407	forward 1	RCC1	Regulator of chromosome condensation	1.60E-04
30	LI:406860.20:2000MAY01	228	407	forward 3	ig	Immunoglobulin domain	1.90E-08
31	LI:142384.1:2000MAY01	318	791	forward 3	UQ_con	Ubiquitin-conjugating enzyme	1.40E-16
32	LI:895427.1:2000MAY01	437	907	forward 2	RhoGAP	RhoGAP domain	1.20E-40
33	LI:757439.1:2000MAY01	1040	1162	forward 2	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	7.20E-10
34	LI:1144066.1:2000MAY01	222	365	forward 3	jmjN	jmjN domain	2.80E-23
35	LI:243660.4:2000MAY01	316	522	forward 1	HMG_box	HMG (high mobility group) box	8.60E-17
36	LI:334386.1:2000MAY01	272	370	forward 2	ank	Ank repeat	4.90E-08
36	LI:334386.1:2000MAY01	735	833	forward 3	ank	Ank repeat	4.50E-05
37	LI:347572.1:2000MAY01	130	1878	forward 1	Peptidase_M2	Angiotensin-converting enzyme	2.60E-05
38	LI:817314.1:2000MAY01	934	2034	forward 1	Trans_recep	Transient receptor	6.50E-260
38	LI:817314.1:2000MAY01	1929	2321	forward 3	Trans_recep	Transient receptor	2.20E-81
39	LI:000290.1:2000MAY01	960	1040	forward 3	zf-CCCH	Zinc finger C-x8-C-x5-C-x3-H type (and s	7.70E-04
40	LI:023518.3:2000MAY01	195	845	forward 3	vATP-synt_AC39	ATP synthase (C/AC39) subunit	5.30E-38
41	LI:1084246.1:2000MAY01	1443	1733	forward 3	cadherin	Cadherin domain	2.30E-20
41	LI:1084246.1:2000MAY01	875	1150	forward 2	cadherin	Cadherin domain	6.60E-17
42	LI:1165828.1:2000MAY01	1421	1705	forward 2	cadherin	Cadherin domain	1.30E-19
43	LI:007302.1:2000MAY01	1646	1810	forward 2	LRRCT	Leucine rich repeat C-terminal domain	2.60E-13
43	LI:007302.1:2000MAY01	1991	2455	forward 2	TIR	TIR domain	3.50E-37
44	LI:236386.4:2000MAY01	677	850	forward 2	SH3	SH3 domain	5.20E-07
45	LI:252904.5:2000MAY01	358	495	forward 1	Kelch	Kelch motif	3.80E-07

TABLE 3

SEQ ID NO:	Template ID	Start	Stop	Frame	Domain Type	Topology
1	LG:977683.1:2000FEB18	373	459	forward 1	TM	N in
1	LG:977683.1:2000FEB18	657	731	forward 3	TM	N out
2	LG:893050.1:2000FEB18	15	101	forward 3	TM	N out
3	LG:980153.1:2000FEB18	313	375	forward 1	TM	N out
3	LG:980153.1:2000FEB18	391	453	forward 1	TM	N out
3	LG:980153.1:2000FEB18	278	364	forward 2	TM	N out
3	LG:980153.1:2000FEB18	416	493	forward 2	TM	N out
3	LG:980153.1:2000FEB18	809	871	forward 2	TM	N out
3	LG:980153.1:2000FEB18	902	964	forward 2	TM	N out
3	LG:980153.1:2000FEB18	1181	1264	forward 2	TM	N out
3	LG:980153.1:2000FEB18	1427	1510	forward 2	TM	N out
3	LG:980153.1:2000FEB18	1733	1798	forward 2	TM	N out
3	LG:980153.1:2000FEB18	1868	1954	forward 2	TM	N out
3	LG:980153.1:2000FEB18	2141	2227	forward 2	TM	N out
3	LG:980153.1:2000FEB18	2261	2308	forward 2	TM	N out
3	LG:980153.1:2000FEB18	60	125	forward 3	TM	N in
3	LG:980153.1:2000FEB18	402	476	forward 3	TM	N in
3	LG:980153.1:2000FEB18	2031	2081	forward 3	TM	N in
3	LG:980153.1:2000FEB18	2142	2213	forward 3	TM	N in
5	LG:475551.1:2000FEB18	2134	2208	forward 1	TM	N in
5	LG:475551.1:2000FEB18	2039	2125	forward 2	TM	N out
5	LG:475551.1:2000FEB18	1167	1217	forward 3	TM	N in
6	LG:481407.2:2000FEB18	874	927	forward 1	TM	
6	LG:481407.2:2000FEB18	949	1035	forward 1	TM	
6	LG:481407.2:2000FEB18	1081	1161	forward 1	TM	
6	LG:481407.2:2000FEB18	1510	1584	forward 1	TM	
6	LG:481407.2:2000FEB18	1355	1435	forward 2	TM	N out
6	LG:481407.2:2000FEB18	1439	1525	forward 2	TM	N out
6	LG:481407.2:2000FEB18	1326	1409	forward 3	TM	N in
6	LG:481407.2:2000FEB18	1446	1526	forward 3	TM	N in
6	LG:481407.2:2000FEB18	1545	1616	forward 3	TM	N in
7	LI:443580.1:2000FEB01	488	574	forward 2	TM	N out
10	LG:171377.1:2000MAY19	318	386	forward 3	TM	N in
10	LG:171377.1:2000MAY19	549	635	forward 3	TM	N in
10	LG:171377.1:2000MAY19	669	740	forward 3	TM	N in
12	LG:247384.1:2000MAY19	1381	1461	forward 1	TM	N in
12	LG:247384.1:2000MAY19	1624	1710	forward 1	TM	N in
12	LG:247384.1:2000MAY19	1409	1495	forward 2	TM	N in
12	LG:247384.1:2000MAY19	1395	1481	forward 3	TM	N in
12	LG:247384.1:2000MAY19	1617	1679	forward 3	TM	N in
13	LG:403872.1:2000MAY19	535	621	forward 1	TM	N in
13	LG:403872.1:2000MAY19	1360	1446	forward 1	TM	N in
13	LG:403872.1:2000MAY19	1522	1581	forward 1	TM	N in
13	LG:403872.1:2000MAY19	1828	1902	forward 1	TM	N in
13	LG:403872.1:2000MAY19	1957	2022	forward 1	TM	N in
13	LG:403872.1:2000MAY19	299	349	forward 2	TM	N in
13	LG:403872.1:2000MAY19	1361	1423	forward 2	TM	N in
13	LG:403872.1:2000MAY19	1439	1501	forward 2	TM	N in
13	LG:403872.1:2000MAY19	1553	1627	forward 2	TM	N in
13	LG:403872.1:2000MAY19	1859	1918	forward 2	TM	N in
13	LG:403872.1:2000MAY19	2027	2110	forward 2	TM	N in
13	LG:403872.1:2000MAY19	2117	2203	forward 2	TM	N in
13	LG:403872.1:2000MAY19	369	452	forward 3	TM	N in

13	LG:403872.1:2000MAY19	549	635	forward 3	TM	N in
13	LG:403872.1:2000MAY19	708	785	forward 3	TM	N in
13	LG:403872.1:2000MAY19	1101	1187	forward 3	TM	N in
13	LG:403872.1:2000MAY19	1419	1505	forward 3	TM	N in
13	LG:403872.1:2000MAY19	1575	1661	forward 3	TM	N in
13	LG:403872.1:2000MAY19	2115	2192	forward 3	TM	N in
13	LG:403872.1:2000MAY19	2226	2273	forward 3	TM	N in
14	LG:1135213.1:2000MAY19	41	127	forward 2	TM	N out
14	LG:1135213.1:2000MAY19	215	274	forward 2	TM	N out
14	LG:1135213.1:2000MAY19	293	379	forward 2	TM	N out
14	LG:1135213.1:2000MAY19	389	475	forward 2	TM	N out
16	LG:342147.1:2000MAY19	142	204	forward 1	TM	N out
16	LG:342147.1:2000MAY19	171	251	forward 3	TM	N out
17	LG:1097300.1:2000MAY19	487	564	forward 1	TM	
17	LG:1097300.1:2000MAY19	805	891	forward 1	TM	
17	LG:1097300.1:2000MAY19	1372	1458	forward 1	TM	
17	LG:1097300.1:2000MAY19	668	754	forward 2	TM	N out
17	LG:1097300.1:2000MAY19	803	874	forward 2	TM	N out
17	LG:1097300.1:2000MAY19	1358	1441	forward 2	TM	N out
17	LG:1097300.1:2000MAY19	522	578	forward 3	TM	N in
17	LG:1097300.1:2000MAY19	750	836	forward 3	TM	N in
17	LG:1097300.1:2000MAY19	894	956	forward 3	TM	N in
17	LG:1097300.1:2000MAY19	1068	1145	forward 3	TM	N in
18	LG:444850.9:2000MAY19	253	315	forward 1	TM	N in
19	LG:402231.6:2000MAY19	407	484	forward 2	TM	N in
23	LG:350793.2:2000MAY19	148	222	forward 1	TM	N in
23	LG:350793.2:2000MAY19	316	384	forward 1	TM	N in
23	LG:350793.2:2000MAY19	1144	1215	forward 1	TM	N in
23	LG:350793.2:2000MAY19	1231	1293	forward 1	TM	N in
23	LG:350793.2:2000MAY19	1339	1425	forward 1	TM	N in
23	LG:350793.2:2000MAY19	1459	1521	forward 1	TM	N in
23	LG:350793.2:2000MAY19	1582	1662	forward 1	TM	N in
23	LG:350793.2:2000MAY19	1882	1953	forward 1	TM	N in
23	LG:350793.2:2000MAY19	1514	1600	forward 2	TM	
23	LG:350793.2:2000MAY19	2135	2221	forward 2	TM	
23	LG:350793.2:2000MAY19	1422	1493	forward 3	TM	
23	LG:350793.2:2000MAY19	2268	2354	forward 3	TM	
24	LG:408751.3:2000MAY19	1202	1264	forward 2	TM	N out
24	LG:408751.3:2000MAY19	1137	1223	forward 3	TM	N in
25	LI:336120.1:2000MAY01	241	297	forward 1	TM	N in
25	LI:336120.1:2000MAY01	616	702	forward 1	TM	N in
25	LI:336120.1:2000MAY01	1141	1200	forward 1	TM	N in
25	LI:336120.1:2000MAY01	2524	2598	forward 1	TM	N in
25	LI:336120.1:2000MAY01	1163	1213	forward 2	TM	N in
25	LI:336120.1:2000MAY01	1922	1972	forward 2	TM	N in
25	LI:336120.1:2000MAY01	2060	2119	forward 2	TM	N in
25	LI:336120.1:2000MAY01	2510	2596	forward 2	TM	N in
25	LI:336120.1:2000MAY01	663	749	forward 3	TM	N in
25	LI:336120.1:2000MAY01	1380	1445	forward 3	TM	N in
25	LI:336120.1:2000MAY01	1839	1925	forward 3	TM	N in
25	LI:336120.1:2000MAY01	2148	2234	forward 3	TM	N in
25	LI:336120.1:2000MAY01	2418	2471	forward 3	TM	N in
25	LI:336120.1:2000MAY01	2499	2585	forward 3	TM	N in
26	LI:234104.2:2000MAY01	1873	1947	forward 1	TM	N out
26	LI:234104.2:2000MAY01	2155	2241	forward 1	TM	N out
26	LI:234104.2:2000MAY01	3616	3690	forward 1	TM	N out

26	LI:234104.2:2000MAY01	1112	1168	forward 2	TM	N in
26	LI:234104.2:2000MAY01	2216	2302	forward 2	TM	N in
26	LI:234104.2:2000MAY01	3632	3718	forward 2	TM	N in
26	LI:234104.2:2000MAY01	3998	4045	forward 2	TM	N in
26	LI:234104.2:2000MAY01	1314	1400	forward 3	TM	N in
26	LI:234104.2:2000MAY01	2172	2258	forward 3	TM	N in
26	LI:234104.2:2000MAY01	2607	2684	forward 3	TM	N in
26	LI:234104.2:2000MAY01	2739	2798	forward 3	TM	N in
26	LI:234104.2:2000MAY01	2841	2891	forward 3	TM	N in
26	LI:234104.2:2000MAY01	3621	3707	forward 3	TM	N in
26	LI:234104.2:2000MAY01	4080	4145	forward 3	TM	N in
28	LI:119992.3:2000MAY01	22	102	forward 1	TM	N out
28	LI:119992.3:2000MAY01	151	237	forward 1	TM	N out
28	LI:119992.3:2000MAY01	1444	1530	forward 1	TM	N out
28	LI:119992.3:2000MAY01	1603	1683	forward 1	TM	N out
28	LI:119992.3:2000MAY01	1729	1809	forward 1	TM	N out
28	LI:119992.3:2000MAY01	2197	2253	forward 1	TM	N out
28	LI:119992.3:2000MAY01	2269	2355	forward 1	TM	N out
28	LI:119992.3:2000MAY01	2989	3075	forward 1	TM	N out
28	LI:119992.3:2000MAY01	3163	3249	forward 1	TM	N out
28	LI:119992.3:2000MAY01	1247	1333	forward 2	TM	N in
28	LI:119992.3:2000MAY01	1538	1606	forward 2	TM	N in
28	LI:119992.3:2000MAY01	2207	2293	forward 2	TM	N in
28	LI:119992.3:2000MAY01	2756	2812	forward 2	TM	N in
28	LI:119992.3:2000MAY01	3098	3169	forward 2	TM	N in
28	LI:119992.3:2000MAY01	3281	3343	forward 2	TM	N in
28	LI:119992.3:2000MAY01	3356	3418	forward 2	TM	N in
28	LI:119992.3:2000MAY01	120	188	forward 3	TM	N in
28	LI:119992.3:2000MAY01	627	689	forward 3	TM	N in
28	LI:119992.3:2000MAY01	708	770	forward 3	TM	N in
28	LI:119992.3:2000MAY01	1425	1511	forward 3	TM	N in
28	LI:119992.3:2000MAY01	1782	1868	forward 3	TM	N in
28	LI:119992.3:2000MAY01	2223	2306	forward 3	TM	N in
28	LI:119992.3:2000MAY01	2757	2843	forward 3	TM	N in
28	LI:119992.3:2000MAY01	3027	3113	forward 3	TM	N in
28	LI:119992.3:2000MAY01	3213	3275	forward 3	TM	N in
28	LI:119992.3:2000MAY01	3312	3374	forward 3	TM	N in
29	LI:197241.2:2000MAY01	289	369	forward 1	TM	N out
29	LI:197241.2:2000MAY01	430	507	forward 1	TM	N out
29	LI:197241.2:2000MAY01	799	861	forward 1	TM	N out
29	LI:197241.2:2000MAY01	889	951	forward 1	TM	N out
29	LI:197241.2:2000MAY01	1798	1863	forward 1	TM	N out
29	LI:197241.2:2000MAY01	1930	2016	forward 1	TM	N out
29	LI:197241.2:2000MAY01	2101	2148	forward 1	TM	N out
29	LI:197241.2:2000MAY01	2206	2262	forward 1	TM	N out
29	LI:197241.2:2000MAY01	416	499	forward 2	TM	N out
29	LI:197241.2:2000MAY01	812	862	forward 2	TM	N out
29	LI:197241.2:2000MAY01	1226	1309	forward 2	TM	N out
29	LI:197241.2:2000MAY01	1475	1558	forward 2	TM	N out
29	LI:197241.2:2000MAY01	2210	2296	forward 2	TM	N out
29	LI:197241.2:2000MAY01	60	125	forward 3	TM	N in
29	LI:197241.2:2000MAY01	333	395	forward 3	TM	N in
29	LI:197241.2:2000MAY01	441	503	forward 3	TM	N in
29	LI:197241.2:2000MAY01	2223	2300	forward 3	TM	N in
31	LI:142384.1:2000MAY01	367	432	forward 1	TM	N out
31	LI:142384.1:2000MAY01	93	155	forward 3	TM	N out

32	LI:895427.1:2000MAY01	1796	1879	forward 2	TM	N in
32	LI:895427.1:2000MAY01	1656	1724	forward 3	TM	N in
33	LI:757439.1:2000MAY01	253	312	forward 1	TM	N in
33	LI:757439.1:2000MAY01	817	900	forward 1	TM	N in
33	LI:757439.1:2000MAY01	1507	1572	forward 1	TM	N in
33	LI:757439.1:2000MAY01	1615	1677	forward 1	TM	N in
33	LI:757439.1:2000MAY01	1696	1758	forward 1	TM	N in
33	LI:757439.1:2000MAY01	1834	1899	forward 1	TM	N in
33	LI:757439.1:2000MAY01	1969	2043	forward 1	TM	N in
33	LI:757439.1:2000MAY01	2107	2193	forward 1	TM	N in
33	LI:757439.1:2000MAY01	2506	2586	forward 1	TM	N in
33	LI:757439.1:2000MAY01	815	901	forward 2	TM	N out
33	LI:757439.1:2000MAY01	1634	1720	forward 2	TM	N out
33	LI:757439.1:2000MAY01	1796	1882	forward 2	TM	N out
33	LI:757439.1:2000MAY01	1952	2026	forward 2	TM	N out
33	LI:757439.1:2000MAY01	2486	2563	forward 2	TM	N out
33	LI:757439.1:2000MAY01	783	869	forward 3	TM	N in
33	LI:757439.1:2000MAY01	996	1049	forward 3	TM	N in
33	LI:757439.1:2000MAY01	1545	1631	forward 3	TM	N in
33	LI:757439.1:2000MAY01	2115	2174	forward 3	TM	N in
35	LI:243660.4:2000MAY01	1247	1333	forward 2	TM	N in
36	LI:334386.1:2000MAY01	538	621	forward 1	TM	
36	LI:334386.1:2000MAY01	922	1008	forward 1	TM	
36	LI:334386.1:2000MAY01	1087	1173	forward 1	TM	
36	LI:334386.1:2000MAY01	1468	1530	forward 1	TM	
36	LI:334386.1:2000MAY01	1570	1632	forward 1	TM	
36	LI:334386.1:2000MAY01	2731	2802	forward 1	TM	
36	LI:334386.1:2000MAY01	2992	3054	forward 1	TM	
36	LI:334386.1:2000MAY01	3325	3387	forward 1	TM	
36	LI:334386.1:2000MAY01	3406	3468	forward 1	TM	
36	LI:334386.1:2000MAY01	3487	3570	forward 1	TM	
36	LI:334386.1:2000MAY01	3766	3852	forward 1	TM	
36	LI:334386.1:2000MAY01	4006	4077	forward 1	TM	
36	LI:334386.1:2000MAY01	4342	4416	forward 1	TM	
36	LI:334386.1:2000MAY01	4615	4686	forward 1	TM	
36	LI:334386.1:2000MAY01	4747	4833	forward 1	TM	
36	LI:334386.1:2000MAY01	5062	5124	forward 1	TM	
36	LI:334386.1:2000MAY01	5140	5202	forward 1	TM	
36	LI:334386.1:2000MAY01	5227	5289	forward 1	TM	
36	LI:334386.1:2000MAY01	5563	5649	forward 1	TM	
36	LI:334386.1:2000MAY01	1235	1321	forward 2	TM	N in
36	LI:334386.1:2000MAY01	2423	2476	forward 2	TM	N in
36	LI:334386.1:2000MAY01	2702	2764	forward 2	TM	N in
36	LI:334386.1:2000MAY01	2792	2854	forward 2	TM	N in
36	LI:334386.1:2000MAY01	3086	3172	forward 2	TM	N in
36	LI:334386.1:2000MAY01	3302	3355	forward 2	TM	N in
36	LI:334386.1:2000MAY01	3452	3517	forward 2	TM	N in
36	LI:334386.1:2000MAY01	3920	4006	forward 2	TM	N in
36	LI:334386.1:2000MAY01	4064	4144	forward 2	TM	N in
36	LI:334386.1:2000MAY01	4250	4318	forward 2	TM	N in
36	LI:334386.1:2000MAY01	4331	4402	forward 2	TM	N in
36	LI:334386.1:2000MAY01	4523	4576	forward 2	TM	N in
36	LI:334386.1:2000MAY01	4586	4669	forward 2	TM	N in
36	LI:334386.1:2000MAY01	4772	4855	forward 2	TM	N in
36	LI:334386.1:2000MAY01	5039	5125	forward 2	TM	N in
36	LI:334386.1:2000MAY01	5498	5584	forward 2	TM	N in

36	LI:334386.1:2000MAY01	30	116	forward 3	TM	N in
36	LI:334386.1:2000MAY01	324	380	forward 3	TM	N in
36	LI:334386.1:2000MAY01	387	470	forward 3	TM	N in
36	LI:334386.1:2000MAY01	531	608	forward 3	TM	N in
36	LI:334386.1:2000MAY01	1362	1448	forward 3	TM	N in
36	LI:334386.1:2000MAY01	1539	1625	forward 3	TM	N in
36	LI:334386.1:2000MAY01	2232	2279	forward 3	TM	N in
36	LI:334386.1:2000MAY01	2580	2651	forward 3	TM	N in
36	LI:334386.1:2000MAY01	2757	2822	forward 3	TM	N in
36	LI:334386.1:2000MAY01	2820	2870	forward 3	TM	N in
36	LI:334386.1:2000MAY01	3282	3368	forward 3	TM	N in
36	LI:334386.1:2000MAY01	3510	3596	forward 3	TM	N in
36	LI:334386.1:2000MAY01	3981	4064	forward 3	TM	N in
36	LI:334386.1:2000MAY01	4356	4427	forward 3	TM	N in
36	LI:334386.1:2000MAY01	4464	4544	forward 3	TM	N in
36	LI:334386.1:2000MAY01	4959	5024	forward 3	TM	N in
36	LI:334386.1:2000MAY01	5601	5687	forward 3	TM	N in
37	LI:347572.1:2000MAY01	790	876	forward 1	TM	N in
37	LI:347572.1:2000MAY01	1354	1434	forward 1	TM	N in
37	LI:347572.1:2000MAY01	2425	2511	forward 1	TM	N in
37	LI:347572.1:2000MAY01	2599	2685	forward 1	TM	N in
37	LI:347572.1:2000MAY01	2686	2757	forward 1	TM	N in
37	LI:347572.1:2000MAY01	3133	3207	forward 1	TM	N in
37	LI:347572.1:2000MAY01	1184	1255	forward 2	TM	
37	LI:347572.1:2000MAY01	2264	2350	forward 2	TM	
37	LI:347572.1:2000MAY01	2597	2665	forward 2	TM	
37	LI:347572.1:2000MAY01	2942	3028	forward 2	TM	
37	LI:347572.1:2000MAY01	3137	3199	forward 2	TM	
37	LI:347572.1:2000MAY01	3227	3289	forward 2	TM	
37	LI:347572.1:2000MAY01	129	215	forward 3	TM	N in
37	LI:347572.1:2000MAY01	969	1046	forward 3	TM	N in
37	LI:347572.1:2000MAY01	1947	2033	forward 3	TM	N in
37	LI:347572.1:2000MAY01	2208	2288	forward 3	TM	N in
37	LI:347572.1:2000MAY01	2412	2477	forward 3	TM	N in
37	LI:347572.1:2000MAY01	2604	2684	forward 3	TM	N in
37	LI:347572.1:2000MAY01	2739	2795	forward 3	TM	N in
38	LI:817314.1:2000MAY01	460	546	forward 1	TM	
38	LI:817314.1:2000MAY01	1192	1278	forward 1	TM	
38	LI:817314.1:2000MAY01	1318	1386	forward 1	TM	
38	LI:817314.1:2000MAY01	1423	1485	forward 1	TM	
38	LI:817314.1:2000MAY01	1537	1599	forward 1	TM	
38	LI:817314.1:2000MAY01	1630	1692	forward 1	TM	
38	LI:817314.1:2000MAY01	1756	1842	forward 1	TM	
38	LI:817314.1:2000MAY01	1930	1992	forward 1	TM	
38	LI:817314.1:2000MAY01	2032	2094	forward 1	TM	
38	LI:817314.1:2000MAY01	2860	2946	forward 1	TM	
38	LI:817314.1:2000MAY01	3127	3213	forward 1	TM	
38	LI:817314.1:2000MAY01	362	448	forward 2	TM	N in
38	LI:817314.1:2000MAY01	3158	3244	forward 2	TM	N in
38	LI:817314.1:2000MAY01	30	95	forward 3	TM	N out
38	LI:817314.1:2000MAY01	1239	1301	forward 3	TM	N out
38	LI:817314.1:2000MAY01	1785	1865	forward 3	TM	N out
38	LI:817314.1:2000MAY01	1920	2000	forward 3	TM	N out
38	LI:817314.1:2000MAY01	3189	3269	forward 3	TM	N out
39	LI:000290.1:2000MAY01	1003	1065	forward 1	TM	N in
39	LI:000290.1:2000MAY01	1075	1137	forward 1	TM	N in

39	LI:000290.1:2000MAY01	1195	1248	forward 1	TM	N in
39	LI:000290.1:2000MAY01	767	844	forward 2	TM	
39	LI:000290.1:2000MAY01	882	932	forward 3	TM	N in
40	LI:023518.3:2000MAY01	28	108	forward 1	TM	N out
40	LI:023518.3:2000MAY01	20	106	forward 2	TM	N in
41	LI:1084246.1:2000MAY01	178	264	forward 1	TM	N out
41	LI:1084246.1:2000MAY01	2686	2760	forward 1	TM	N out
41	LI:1084246.1:2000MAY01	2932	3003	forward 1	TM	N out
41	LI:1084246.1:2000MAY01	3097	3159	forward 1	TM	N out
41	LI:1084246.1:2000MAY01	3184	3246	forward 1	TM	N out
41	LI:1084246.1:2000MAY01	3352	3405	forward 1	TM	N out
41	LI:1084246.1:2000MAY01	3409	3480	forward 1	TM	N out
41	LI:1084246.1:2000MAY01	3526	3609	forward 1	TM	N out
41	LI:1084246.1:2000MAY01	200	253	forward 2	TM	N in
41	LI:1084246.1:2000MAY01	2171	2254	forward 2	TM	N in
41	LI:1084246.1:2000MAY01	2654	2734	forward 2	TM	N in
41	LI:1084246.1:2000MAY01	3065	3142	forward 2	TM	N in
41	LI:1084246.1:2000MAY01	3284	3358	forward 2	TM	N in
41	LI:1084246.1:2000MAY01	3479	3553	forward 2	TM	N in
41	LI:1084246.1:2000MAY01	582	641	forward 3	TM	N out
41	LI:1084246.1:2000MAY01	2127	2213	forward 3	TM	N out
41	LI:1084246.1:2000MAY01	2457	2543	forward 3	TM	N out
41	LI:1084246.1:2000MAY01	2580	2666	forward 3	TM	N out
41	LI:1084246.1:2000MAY01	2751	2813	forward 3	TM	N out
41	LI:1084246.1:2000MAY01	2826	2888	forward 3	TM	N out
41	LI:1084246.1:2000MAY01	2961	3047	forward 3	TM	N out
41	LI:1084246.1:2000MAY01	3249	3335	forward 3	TM	N out
41	LI:1084246.1:2000MAY01	3429	3515	forward 3	TM	N out
42	LI:1165828.1:2000MAY01	61	147	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	244	312	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	454	510	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	3664	3750	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	3937	4023	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	4600	4653	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	4855	4941	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	5047	5133	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	5227	5298	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	5311	5388	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	5491	5577	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	5800	5871	forward 1	TM	N out
42	LI:1165828.1:2000MAY01	227	301	forward 2	TM	N in
42	LI:1165828.1:2000MAY01	713	775	forward 2	TM	N in
42	LI:1165828.1:2000MAY01	1769	1819	forward 2	TM	N in
42	LI:1165828.1:2000MAY01	2759	2845	forward 2	TM	N in
42	LI:1165828.1:2000MAY01	3869	3928	forward 2	TM	N in
42	LI:1165828.1:2000MAY01	4688	4774	forward 2	TM	N in
42	LI:1165828.1:2000MAY01	5048	5116	forward 2	TM	N in
42	LI:1165828.1:2000MAY01	5531	5617	forward 2	TM	N in
42	LI:1165828.1:2000MAY01	5816	5893	forward 2	TM	N in
42	LI:1165828.1:2000MAY01	39	113	forward 3	TM	N out
42	LI:1165828.1:2000MAY01	906	968	forward 3	TM	N out
42	LI:1165828.1:2000MAY01	1602	1688	forward 3	TM	N out
42	LI:1165828.1:2000MAY01	3471	3557	forward 3	TM	N out
42	LI:1165828.1:2000MAY01	3558	3608	forward 3	TM	N out
42	LI:1165828.1:2000MAY01	4203	4289	forward 3	TM	N out
42	LI:1165828.1:2000MAY01	4749	4835	forward 3	TM	N out

42	LI:1165828.1:2000MAY01	5625	5690	forward 3	TM	N out
42	LI:1165828.1:2000MAY01	5847	5918	forward 3	TM	N out
43	LI:007302.1:2000MAY01	346	426	forward 1	TM	N in
43	LI:007302.1:2000MAY01	2638	2721	forward 1	TM	N in
43	LI:007302.1:2000MAY01	59	145	forward 2	TM	N out
43	LI:007302.1:2000MAY01	653	718	forward 2	TM	N out
43	LI:007302.1:2000MAY01	1799	1885	forward 2	TM	N out
43	LI:007302.1:2000MAY01	321	407	forward 3	TM	N in
43	LI:007302.1:2000MAY01	480	566	forward 3	TM	N in
43	LI:007302.1:2000MAY01	645	704	forward 3	TM	N in
43	LI:007302.1:2000MAY01	807	890	forward 3	TM	N in
43	LI:007302.1:2000MAY01	1161	1223	forward 3	TM	N in
43	LI:007302.1:2000MAY01	1236	1298	forward 3	TM	N in
43	LI:007302.1:2000MAY01	1362	1448	forward 3	TM	N in
43	LI:007302.1:2000MAY01	1809	1868	forward 3	TM	N in
43	LI:007302.1:2000MAY01	1998	2084	forward 3	TM	N in
43	LI:007302.1:2000MAY01	2184	2234	forward 3	TM	N in
43	LI:007302.1:2000MAY01	2457	2540	forward 3	TM	N in
43	LI:007302.1:2000MAY01	2595	2681	forward 3	TM	N in
44	LI:236386.4:2000MAY01	3739	3792	forward 1	TM	N out
44	LI:236386.4:2000MAY01	53	118	forward 2	TM	N out
44	LI:236386.4:2000MAY01	218	304	forward 2	TM	N out
44	LI:236386.4:2000MAY01	3755	3823	forward 2	TM	N out
44	LI:236386.4:2000MAY01	2376	2435	forward 3	TM	N out
45	LI:252904.5:2000MAY01	494	550	forward 2	TM	N out
45	LI:252904.5:2000MAY01	300	374	forward 3	TM	N out

TABLE 4

SEQ ID NO:	Component ID	Start	Stop
1	g5813583	610	959
1	6817504J1	1	621
1	g1989978	3	264
1	4292280H1	10	242
1	483000R6	11	337
1	483000H1	11	252
1	g1424329	14	316
1	3255214H1	107	349
1	1450061H1	131	371
1	5388816H1	152	419
1	955673H1	181	406
1	2109273H1	286	547
1	5980116H1	373	651
1	g828864	376	596
1	3072657H1	380	488
1	2949928H1	416	680
1	6016294H1	580	677
1	g1855323	611	695
1	g1623907	611	667
1	g1855498	611	933
1	g1751162	689	928
1	1309114T6	716	955
1	1309114F6	716	979
1	1309114H1	716	971
1	3637614H1	807	1053
1	7065033H1	899	1165
1	6817504H1	971	1358
1	6013754H1	978	1245
1	g573231	1034	1316
1	g709283	1034	1322
1	g767017	1035	1345
1	g692230	1061	1388
1	1617090H1	1084	1209
1	1617090F6	1084	1380
1	g1157664	1112	1412
2	6131346H1	1	193
2	6871387H1	125	662
2	g2279352	352	634
3	7039759H1	1390	1914
3	6481201H1	1428	1542
3	6929893H1	1460	1891
3	160750H1	1643	1734
3	6201684H1	1659	2172
3	492554H1	36	275
3	6710369H1	84	594
3	g770845	369	639
3	6710369J1	538	1037
3	6866894H1	749	1339
3	2045879F6	796	1123
3	2045879H1	796	1064
3	g677645	854	1153
3	g570913	854	1235
3	2837088H1	1	79
3	g878213	855	1194
3	3637810H1	905	1188
3	382301R6	11	244
3	3637810F8	906	1347
3	5516287H1	938	1192
3	382301H1	11	273
3	310657H1	983	1184
3	381716R1	11	471
3	054856H1	1027	1268
3	2676843H1	1102	1294
3	2865460H1	1182	1413
3	5983503H1	1223	1521
3	3296833H1	24	289
3	492559R1	36	564
3	3903656H1	1288	1501
3	2554026H1	1322	1591
3	g1894266	1326	1800
3	3151953H1	2028	2266
3	6357422H1	2056	2344
3	382301T6	2063	2619
3	2498615F6	2077	2500
3	2498615H1	2077	2310
3	492559F1	2104	2658
3	2684917H1	1709	1950
3	3898190H1	1917	2210
3	381716F1	2106	2658
3	5952437H1	1960	2247
3	4701147H1	2134	2402
3	g5435909	2213	2663
3	7067611H1	2254	2764
3	g2563607	2282	2658
3	1889064H1	2300	2577
3	2400488H1	2302	2549
3	g817549	2307	2667
3	g566965	2343	2658
3	g1894154	2354	2658
3	g869609	2394	2667
3	g4291206	2396	2766
3	g646309	2398	2658
3	3249908H1	2467	2760
3	672907H1	2516	2658
3	672763R6	2516	2658
3	672763H1	2516	2658
3	672696H1	2516	2658
3	672763T6	2516	2621
4	g1939101	219	609

TABLE 4 (cont.)

4	1749048T6	1	388	5	1515410H1	1224	1442	5	4671595H1	2027	2277		
5	996489H1	1	289	5	g2056082	1221	1509	5	318659H1	2041	2291		
5	996489R6	1	321	5	566614H1	1269	1530	5	4902185H1	2096	2297		
5	6807726H1	9	414	5	4780315H1	1290	1553	5	g2055975	2105	2298		
5	g1208184	74	603	5	1637781H1	1302	1454	5	1219763H1	2110	2288		
5	g1146490	110	406	5	1638827H1	1302	1455	5	1219763R6	2110	2290		
5	1391557H1	145	273	5	1633937H1	1762	1969	5	1219763T6	2110	2251		
5	2054016H1	155	406	5	6821354H1	1419	1971	5	1219763T1	2110	2250		
5	3564377H1	213	498	5	1390745H1	1433	1557	5	581809H1	2110	2369		
5	1389469H1	365	607	5	1932110H1	1712	1868	5	g2788727	2119	2369		
5	6178475H1	288	554	5	1932110F6	1713	1960	5	2753294H1	2255	2364		
5	2490333H1	461	684	5	1850028H1	1728	1970	6	2055577R6	766	1137		
5	1498011F6	497	816	5	386578H1	1753	2029	6	2055577T6	766	1096		
5	1498011H1	497	735	5	1862471H1	1759	1870	6	g1578280	767	1137		
5	154577H1	512	727	5	458296H1	1799	1890	6	g4897043	769	1147		
5	2439861H1	600	846	5	2028756H1	1816	1890	6	g1897641	769	1137		
5	6974170H1	655	1206	5	1988349T6	1824	2253	6	g3004281	774	1138		
5	5557446H1	723	990	5	1498011T6	1829	2254	6	6361438H2	776	1335		
5	6821354J1	725	1336	5	6157225H1	1842	2101	6	1273945F1	790	1131		
5	3801324H1	751	1035	5	521110H1	1850	1975	6	1273945H1	790	948		
5	159257H1	753	952	5	6157733H1	1854	2051	6	2558966H1	791	1058		
5	1562163H1	801	1030	5	4829815H1	1889	1962	6	g2178992	831	1147		
5	7161127H1	827	1358	5	4411517H1	1907	2157	6	g1891843	842	1143		
5	1840238H1	834	989	5	541981H1	1927	2155	6	g1203333	844	1159		
5	1892815H1	944	1194	5	4558860H1	1944	2106	6	g1141073	845	1135		
5	1893046H1	944	1185	5	1391452T6	1958	2260	6	g1728655	851	1143		
5	1391452H1	962	1131	5	2752758H1	1963	2239	6	4618322H1	860	1133		
5	1391452F6	962	1223	5	1807380T6	1965	2250	6	g3179203	882	1147		
5	1680496H1	1117	1345	5	1807042F6	1970	2290	6	4164817H1	9	261		
5	2132470R6	1120	1456	5	1807042H1	1970	2255	6	5851107H1	12	270		
5	1265470H1	1149	1401	5	2311115H1	1992	2237	6	4938618H1	1	285		
5	6804038H1	1164	1555	5	996489T6	1994	2332	6	2096384H1	13	274		
5	3430883H1	1183	1428	5	6125387H1	2007	2356	6	4938518H1	1	184		
5	2132470H1	1188	1456	5	4905520H1	2022	2280	6	6133436H1	6	304		

TABLE 4 (cont.)

6	5218795H1	14	282	6	768284H1	670	900	6	5346772H1	29	227
6	3038155H1	6	294	6	g2567185	671	1075	6	5346890H1	29	141
6	3088308H1	14	285	6	2522538H1	672	909	6	4151612H1	31	258
6	6821608H1	14	578	6	g3446544	676	1136	6	g2229063	27	371
6	5855412H1	14	297	6	4377572H1	680	948	6	3074071H1	31	308
6	2532161H1	6	258	6	g4242762	685	1135	6	3717427H1	32	401
6	5999068H1	6	559	6	g5444329	685	1147	6	2467222H1	32	258
6	g5431297	7	324	6	g4394905	687	1135	6	5687205H1	33	296
6	2715577H1	14	256	6	g4891466	689	1136	6	g2027890	31	188
6	3717266H1	6	312	6	4534880T1	604	1111	6	2864630H1	34	341
6	3088671H1	14	251	6	g1422487	626	919	6	3837823H1	35	321
6	1690850T6	16	558	6	3213475H1	692	929	6	5978027H1	35	298
6	4978332H1	19	305	6	g3674532	698	1150	6	3841249H1	35	236
6	2525160H1	368	619	6	g3665343	700	1135	6	5780416H1	37	313
6	2811816H1	382	591	6	g5365390	705	1135	6	4525495H1	38	294
6	5285481H1	381	530	6	3362353H1	708	848	6	2943180H1	35	281
6	g1923667	380	575	6	g3737258	707	1140	6	3159688H1	36	136
6	2724519H1	385	586	6	3801387H1	711	869	6	g2156554	35	459
6	4403213H1	397	537	6	g1277444	717	1135	6	5989823H1	38	334
6	2525196H1	368	597	6	6045963H1	722	1176	6	4525695H1	38	287
6	g2111237	370	592	6	g2236500	716	1139	6	774424H1	38	269
6	g1155753	370	731	6	4024228H1	722	1008	6	4376239H1	38	242
6	g2111348	371	598	6	g4088002	718	1149	6	222536R1	19	533
6	g3798474	371	588	6	3553263H1	754	969	6	4951501H2	19	325
6	g2968466	372	670	6	g2229274	762	1153	6	5986222H1	21	289
6	g1874430	374	675	6	2055577H1	766	1031	6	4782312H1	19	258
6	g3933996	376	589	6	5116334H1	19	290	6	222536H1	19	150
6	g2567131	409	663	6	1546662H1	19	218	6	6152094H1	26	301
6	g1422584	429	556	6	2275605H1	19	291	6	3365655H1	27	286
6	g2157052	435	744	6	5968841H1	19	591	6	2098005H1	27	209
6	3092788H1	437	722	6	1902261H1	1	288	6	2874828H1	27	311
6	1650634F6	441	871	6	6728620H1	29	590	6	4748012H1	29	297
6	1831391H1	637	867	6	1690850F6	29	482	6	5122477H1	27	278
6	2173245H1	652	888	6	1690850H1	29	237	6	5516387H1	27	270

TABLE 4 (cont.)

6	5695974H1	27	203	6	5609131H1	123	365	6	g5849856	504	739		
6	4994832H1	36	185	6	g3598018	135	590	6	6365612H1	519	816		
6	g1728758	40	325	6	g3432506	136	593	6	5183801H1	525	789		
6	5993725H1	40	342	6	g5431490	144	323	6	3706413H1	529	812		
6	5995510H1	40	330	6	g1646810 -	57	324	6	4828553H1	532	762		
6	g4329715	40	406	6	g2555607	156	500	6	2604912H1	539	791		
6	2894305H1	47	310	6	g1578371	53	198	6	g2107086	553	977		
6	2719394T6	303	625	6	g2229126	158	593	6	g5769539	555	733		
6	g5658221	327	736	6	g3229125	173	598	6	5576107H1	559	800		
6	5857676H1	296	564	6	g3898868	173	593	6	g1891969	565	972		
6	5726056H2	297	676	6	g4452177	180	323	6	3620132H1	31	324		
6	2097760H1	300	546	6	g3182012	205	593	6	4605074H1	598	846		
6	2873090H1	329	605	6	790141R1	222	746	6	1650642F6	441	832		
6	3136434H1	334	597	6	790141H1	222	456	6	3443641H1	484	742		
6	g1646811	339	596	6	3599189H1	229	519	6	g3889543	490	917		
6	2738075F6	321	767	6	g2204943	229	593	6	g3095491	492	586		
6	2738075H1	321	564	6	3258218H1	232	529	6	2738075T6	494	1096		
6	2719394F6	318	683	6	g2355330	244	592	6	4534880H1	441	701		
6	2719394H1	267	521	6	g2882852	65	382	6	4277322H1	497	751		
6	g5527461	339	586	6	g1950563	70	330	6	4989476F8	496	967		
6	g2437242	340	551	6	1548020H1	72	301	6	1650634H1	441	687		
6	4724150H1	343	607	6	2823270H1	250	538	6	g2575167	443	843		
6	g1312816	346	778	6	2873603H1	257	537	6	3718361H1	456	769		
6	4787470H1	360	597	6	2755517H1	79	346	6	3267371H1	457	700		
6	5003922H1	362	616	6	3718262H1	81	391	6	1902161H1	462	586		
6	6156796H1	87	345	6	915491R6	260	597	6	5056004H1	465	746		
6	2895320H1	43	273	6	915491H1	260	569	6	g3751871	477	736		
6	4665825H1	96	339	6	4979613H1	276	550	6	2997314H1	482	786		
6	3232485H1	44	316	6	6821608J1	278	791	6	2996840H1	483	745		
6	2399837H1	98	322	6	3246153H1	278	516	6	4276994H1	497	635		
6	6904948H1	101	462	6	4008733H1	281	559	6	g1923480	981	1130		
6	6411519H1	45	554	6	4989076H1	497	752	6	6550669H1	1020	1619		
6	035304H1	55	324	6	g5850851	503	739	6	g4083790	1388	1829		
6	4573015H1	116	388	6	g4738819	504	739	6	4700302H1	1388	1666		

TABLE 4 (cont.)

6	g3770915	1402	1832
6	g1224283	1032	1442
6	g2767747	1055	1135
6	2539090H1	1087	1334
6	1773532H1	1179	1391
6	6045963J1	1211	1801
6	1650634T6	1270	1789
6	g4373516	1308	1756
7	g2524924	315	730
7	g2161228	313	724
7	g3802198	329	703
7	g3147794	231	688
7	g2162211	119	550
7	2497157H1	78	310
7	2854513H1	1	290
8	1985316H1	1	269
8	1985316R6	1	310
8	197972T6	43	445
8	197972H1	43	274
8	197972R6	43	457
9	7197754H2	1	582
10	g5810426	1	449
10	g2219401	2	423
10	g4329377	27	489
10	g2537784	172	669
10	g1376965	259	669
10	4983705H1	270	539
10	7269840H1	339	848
11	6453567H1	1	503
11	4052122H1	185	457
11	4052122F7	185	636
11	g3897399	255	371
12	973628H1	996	1226
12	3014231H1	1097	1369
12	975169T6	1112	1714
12	3042767T6	1122	1713
12	6218188H1	1165	1678
12	5151940H1	1216	1440
12	975304T6	1231	1709
12	5531975T6	1266	1741
12	3577265H1	1286	1598
12	3016255H1	1291	1599
12	970343R6	1304	1757
12	970343H1	1304	1606
12	970343T6	1322	1714
12	3575519H1	1334	1616
12	5153116H1	1345	1469
12	988837H1	1422	1684
12	g4088627	1503	1756
12	6903302H1	1564	2110
12	975169H1	856	1057
12	g2156118	1	475
12	975304H1	2	248
12	3403717H1	1	249
12	4042617H1	1	256
12	3042767H1	3	267
12	3042767F6	3	275
12	4854092H1	4	234
12	4743545H1	6	265
12	5856186H1	20	270
12	535036H1	27	246
12	3960535H2	379	641
12	3960535F6	379	742
12	6216170H1	579	726
12	4456047H1	621	886
12	945050H1	762	1003
12	920681H1	855	1174
12	923436H1	855	1167
12	975169R6	855	1336
13	4745248H1	1	241
13	7158869H1	7	479
13	3335250F6	34	398
13	3335250H1	34	273
13	7077668H1	136	659
13	4318873H1	159	370
13	6992614H1	236	740
13	753174H1	356	543
13	7046749H1	453	1036
13	6983112H1	621	891
13	g570318	630	905
13	5266308H1	632	788
13	g778569	673	993
13	748982H1	672	901
13	744829R1	672	1226
13	744829H1	672	902
13	g869715	672	1004
13	g565684	901	1080
13	g1025621	1027	1340
13	g1059514	1027	1251
13	g714830	1108	1397
13	4311224H1	1203	1484
13	2292254R6	1398	1866
13	2292421R6	1398	1506
13	2291932H1	1398	1649
13	530715H1	1423	1644
13	7090888H1	1520	1659
13	g3086021	1518	1916
13	2291932T6	1559	2132
13	3335250T6	1562	2050
13	6841962H1	1748	2279
13	6855669H1	1881	2375
13	746910R6	1912	2375

TABLE 4 (cont.)

13	746910H1	1912	2143
13	746910T6	1913	2371
13	6844175H1	1941	2375
13	2568562H1	1989	2222
13	g4393425	1996	2415
13	g4109519	2006	2375
13	g2694947	2036	2375
13	g2703845	2040	2375
13	g3884077	2042	2375
13	g3278030	2045	2423
13	4705947H1	2104	2256
13	g714831	2110	2411
13	750787H1	2121	2365
13	667235H1	2126	2370
13	g561290	2150	2375
13	g518739	2157	2375
13	g3230679	2187	2375
13	g717890	2318	2390
14	4145560H1	1	337
14	7182979H1	1	537
14	g4929686	1	1581
14	g1881193	113	359
14	798770H1	206	449
14	g1198695	214	498
14	g1637735	380	642
14	g2204679	39	511
14	5540595H1	1	195
14	g1970769	1	345
14	g1970753	1	325
14	g1971048	1	253
14	g1970777	1	223
14	g815792	8	284
14	g1441646	3	303
14	g4372035	14	479
14	g2930515	35	487
14	g4897951	44	477
14	609028H1	27	178
14	g2782816	15	417
14	g4326525	1	141
14	g2525795	28	236
15	g6450570	1077	1426
15	g6473965	97	472
15	525308H1	117	324
15	g2898932	121	456
15	526619H1	129	370
15	g2942591	134	271
15	2360586H1	145	399
15	2211028H1	228	438
15	987239R1	305	763
15	987239H1	305	478
15	1436565F1	354	824
15	7161757H1	1	521
15	g4372435	23	212
15	g5451540	23	516
15	g3884494	40	407
15	g5545276	40	499
15	2269559H1	44	305
15	2269559R6	44	350
15	g5152652	62	224
15	3222733H1	86	303
15	1664718F6	91	349
15	1664718H1	91	352
15	g880746	97	278
15	1436565H1	354	626
15	2520441H1	360	641
15	3460138H1	393	644
15	6881873J1	142	680
15	6881873H1	51	484
15	1670270F6	637	1077
15	g1921208	645	985
15	6523810H1	659	1052
15	3499282H1	423	706
15	5852917H1	661	921
15	2247228H1	692	959
15	g851799	704	1030
15	4946358H1	711	972
15	5951390H1	729	954
15	6345162H1	792	1031
15	3436737H1	794	1029
15	g2264229	426	815
15	3496822H1	430	703
15	6321740H1	805	1031
15	2112334H1	820	1080
15	1007012H1	470	767
15	2112334R6	820	1167
15	3215530H1	491	714
15	3144904H1	873	1217
15	g4073140	965	1444
15	g4523268	970	1426
15	g5673767	972	1444
15	2836020H1	496	741
15	960106H1	971	1049
15	962045H1	971	1248
15	5109444H1	498	723
15	g2070246	973	1335
15	g2206523	973	1266
15	g880857	501	815
15	g5637498	978	1401
15	g5449171	979	1439
15	3733518H1	980	1275
15	g4763832	981	1444
15	6807693H1	520	1140

TABLE 4 (cont.)

15	1968707R6	522	920
15	g5754504	985	1444
15	g5511006	992	1444
15	6154958H1	991	1304
15	g2952676	993	1443
15	1968707H1	522	727
15	961381H1	997	1290
15	6344762H1	534	632
15	959580H1	997	1109
15	g2209838	548	972
15	6856259H1	554	1067
15	2479125H1	565	804
15	4345262H1	577	856
15	959580R1	997	1433
15	g4437873	998	1426
15	g5661623	1002	1410
15	g4332091	1006	1444
15	5031758H1	585	825
15	g1320158	1008	1439
15	g5391778	1012	1444
15	g5933236	1012	1444
15	g2901335	1014	1408
15	g1940416	1015	1444
15	g5113563	1021	1444
15	2517547H1	1043	1277
15	g5451354	1053	1284
15	g2220466	1062	1408
15	g2952784	1064	1440
15	3329431H1	607	885
15	5271370H1	618	855
15	1670270H1	637	862
15	g1367649	1071	1444
15	g3751105	1073	1444
15	g1367704	1083	1437
15	g5904784	1090	1444
15	g4852367	1094	1444
15	g1443408	1101	1445
15	2124915H1	1117	1402
15	g3412275	1126	1443
15	g5671642	1138	1407
15	g2056619	1211	1442
15	g4148637	1249	1426
15	g1921308	1253	1445
15	g2952936	1256	1443
15	g2728303	1276	1446
15	g4195307	1314	1444
15	g2841540	1351	1445
16	1601184H1	304	515
16	3540611H1	297	388
16	3111986H1	304	368
16	1673924H1	297	503
16	1569636H1	297	508
16	2696549F6	297	378
16	g2219716	1	359
16	g2898608	1	211
16	6755069H1	1	654
16	3539560H1	303	476
16	1515102H1	297	466
16	1572728H1	297	492
16	1347783H1	309	435
16	1691349H1	297	436
16	3686316H1	304	498
17	4563458H1	1	197
17	4381069H1	15	261
17	6205262H1	107	542
17	6202507H1	412	921
17	4620133F6	603	940
17	4620133H1	603	851
17	2158854T6	743	1154
17	g5543295	743	1201
17	g1385006	749	1056
17	2158854H1	749	1012
17	3973473H1	782	1055
17	3973473F8	783	1307
17	5629236F6	806	1288
17	3973473T8	883	1519
17	5629236H1	1062	1288
17	2777742H1	1069	1170
17	2509368H1	1108	1343
17	2793074H2	1138	1253
17	2793074F6	1142	1253
17	2793074T6	1177	1260
17	2364001H1	1404	1651
17	g3898774	1582	1927
18	3224948H1	1	177
18	3695977H1	7	312
18	7006140H1	8	566
18	2794410H1	13	150
18	6460326H1	40	396
18	6787346H1	51	555
18	3403667H1	53	289
18	3725949H1	56	297
18	2830626H1	61	333
18	g1646403	62	445
18	2830626F6	61	581
18	6784569H2	61	591
18	5959276H1	74	534
18	6804522J1	100	522
18	3697994H1	118	356
18	581170H1	133	223
18	5610623H1	133	408
18	2770068H1	157	405

TABLE 4 (cont.)

18	7165406H1	159	535
18	6702265H1	312	825
18	7037116H1	372	699
18	6531787H1	511	922
18	1214116H1	519	662
18	6804522H1	637	1171
18	7218713H1	677	1237
18	3557937H1	687	987
18	6455665H1	825	1420
18	6701662H1	821	1297
18	6523244H1	847	1324
18	4004887H1	926	1204
18	4876106H1	945	1182
18	4067628F7	1082	1353
18	6932868H1	1082	1543
18	3191237H1	1103	1414
18	7088151H1	1126	1596
18	2818868H1	1173	1275
18	5582555H1	1189	1439
18	5582587H1	1188	1442
18	g4893540	1220	1631
18	4442851H1	1276	1544
18	3022715H1	1325	1618
18	3780205H1	1349	1644
18	g1947313	1365	1595
18	2996242H1	1384	1678
18	3052021H1	1414	1704
18	g3095711	1478	1951
18	3927236H1	1596	1856
18	2769806H1	1625	1854
18	5866616H1	1749	1842
18	3730361H1	1767	1870
18	7169445H1	1	343
19	6546889H1	1	339
19	1651460H1	83	301
19	6264819H1	186	461
19	4753777H1	214	338
19	2331424R6	333	638
19	2331424H1	333	560
19	3398569H1	339	582
19	2435387H1	342	570
19	506031H1	351	527
19	6118353H1	362	469
19	609565H1	377	628
19	2873416H1	397	540
20	2583409H1	204	430
20	g2823866	1	383
20	3488619H1	1	280
20	5633561F6	207	798
20	5633561H1	207	465
21	g4690049	1	195
21	1398471F6	1	410
21	1399832H1	1	227
21	2694772H1	126	337
21	2694772F6	125	338
21	1398471H1	1	238
22	5286647F9	1	615
22	3808866F8	5	457
22	7264977H1	17	605
22	4760775F6	38	607
22	5286647T9	242	819
22	5286647T8	506	825
22	5286647F8	5	552
23	628206T7	1954	2472
23	277808H1	1974	2264
23	278730H1	1976	2309
23	275057H1	1976	2160
23	275257H1	1976	2193
23	2586194T6	1977	2477
23	6479875H1	1990	2477
23	2856722H1	2000	2267
23	1298131T6	2038	2472
23	1298131H1	2038	2291
23	1298131F1	2038	2276
23	1298131F6	2038	2516
23	2300965T6	2040	2476
23	g4075934	2067	2517
23	g3415730	2098	2518
23	g2139392	2111	2489
23	g4735514	2111	2514
23	g4261130	2111	2518
23	g4665764	2111	2513
23	3483466H1	2111	2363
23	g5366013	2119	2512
23	g4599402	2126	2517
23	4096757H1	2144	2441
23	2254547H1	2151	2380
23	g1692867	2185	2513
23	g1157366	2204	2513
23	g1128313	2281	2514
23	g2524394	2295	2514
23	g1227222	2316	2513
23	5913552F6	2405	2537
23	2265479H1	2413	2516
23	5913552H1	2416	2504
23	5643316H1	1884	2089
23	5794438H1	1854	2089
23	5791230H1	1854	2089
23	5791375H1	1854	2089
23	856338H1	1129	1361
23	3280567H1	1148	1399
23	6551617H1	1183	1732

TABLE 4 (cont.)

23	6552317H1	1183	1762
23	6751972H1	1191	1762
23	5759260H1	1193	1468
23	4190084H1	1198	1471
23	6136366H1	1270	1571
23	4205570H1	1301	1533
23	3354295H1	1305	1539
23	4303867H1	1317	1502
23	628206H1	1382	1615
23	628206R7	1382	1793
23	4337705H1	1443	1782
23	2881556H1	1467	1726
23	6875744H1	1469	2058
23	5677351H1	1496	1741
23	2772870H1	1505	1749
23	1212235R6	1541	1990
23	1212235H1	1541	1815
23	g1646733	1551	1869
23	2297674H2	1562	1829
23	2586194H1	1590	1839
23	2586194F6	1590	2059
23	2403715H1	1606	1845
23	6859287H1	1655	2089
23	5091604H1	1689	1969
23	2736946H1	1689	1940
23	2823882H1	1714	2005
23	2821225H1	1714	2025
23	573737H1	1740	1857
23	6350742H1	1769	2058
23	2300965H1	1775	2006
23	2300965R6	1775	2170
23	439474H1	1808	2043
23	5686929H1	1843	2106
23	5794171H1	1854	2162
23	5792646H1	1854	2162
23	5792285H1	1854	2089
23	5793871H1	1854	2089
23	4358460H1	1059	1303
23	g2142328	1	284
23	5662770H1	1	178
23	7004664H1	142	653
23	g1692967	194	528
23	265733H1	224	448
23	6406758H1	542	995
23	6259622H1	667	954
23	g1628822	753	1138
23	2587028H1	876	1152
23	3331574H1	913	1177
23	705890H1	915	1149
23	705979H1	915	1181
23	4114902H1	922	1125
23	2889650H1	968	1241
23	6507226H1	1058	1499
23	6258095H1	1059	1340
24	g314920	1324	1655
24	g615297	1324	1656
24	g517687	1324	1655
24	g615578	1370	1656
24	g614283	1374	1656
24	1456735T6	1422	1622
24	g4328099	1446	1662
24	g614262	1449	1656
24	g4152278	1455	1656
24	g562532	1461	1656
24	g671207	1462	1656
24	5945223H1	1578	1660
24	g2985356	1621	1848
24	5498383R6	1236	1619
24	4717574T6	1186	1635
24	1476570F6	1188	1656
24	1476571F6	1188	1532
24	1476570H1	1188	1394
24	g614326	1200	1657
24	1476571T6	1206	1619
24	g4152280	1219	1388
24	g4598685	1229	1657
24	g314775	1244	1656
24	2153570H1	1241	1515
24	4492503H1	1247	1657
24	g615988	1254	1656
24	g775420	1264	1670
24	5659105H1	1264	1344
24	g4617815	1272	1663
24	g5511164	1274	1656
24	g3649444	1275	1658
24	g314750	1287	1656
24	004952H1	1164	1423
24	1476570T6	1171	1617
24	4705993T9	1106	1554
24	1270695T6	1177	1617
24	2416693T6	1090	1611
24	748579R1	1076	1656
24	859218H1	1007	1221
24	g6086997	903	1254
24	533539T6	909	1226
24	5371992T9	942	1580
24	g314842	948	1254
24	g683067	970	1254
24	7290682H1	978	1513
24	009349H1	761	1103
24	6888770H1	772	1287
24	6866213H1	772	1377

TABLE 4 (cont.)

24	4943311T6	785	1231
24	7292792H1	793	1368
24	g1192539	802	1254
24	g4223790	815	1254
24	6717166H1	821	1283
24	g3331126	836	1256
24	5310872H1	838	1064
24	5267191H1	858	1118
24	4940779H1	878	1150
24	1270258H1	880	1118
24	g794503	887	1267
24	g816007	884	1243
24	g901436	892	1254
24	6869327H1	724	1228
24	6855475H1	1045	1242
24	1270292T6	1048	1610
24	g822109	1058	1267
24	748579H1	1064	1304
24	859218R6	1007	1446
24	g567610	1012	1254
24	859218R1	1007	1527
24	859218T6	1046	1617
24	1270695F6	541	829
24	1270695H1	541	773
24	7067123H1	525	1069
24	6448066H1	400	951
24	g691925	443	755
24	533539R6	431	951
24	533539H1	427	622
24	5379139H1	434	679
24	6868778H1	494	1123
24	5674272H1	391	645
24	6120160H1	386	785
24	6866026H1	381	974
24	1456735F6	189	605
24	6721132H1	193	579
24	4203426H1	212	337
24	1992224H1	206	475
24	7259028H1	204	579
24	g766593	289	587
24	7058996H1	305	886
24	4092963H1	327	609
24	g614162	336	605
24	g677813	336	565
24	6985794H1	332	788
24	4338771H1	359	628
24	g708822	393	694
24	g764692	395	736
24	g816062	378	790
24	3864471H1	374	591
24	6990907H1	383	921
24	g1627181	208	330
24	5311056H1	591	753
24	5907142H1	659	938
24	5924427H1	681	971
24	2707020H1	557	850
24	5205391H1	565	805
24	5498383H1	573	811
24	5498383F6	573	1055
24	g4152281	207	277
24	7290347H1	188	672
24	1265660F1	176	785
24	1265660H1	181	469
24	3944530H1	184	461
24	g677040	204	322
24	g1950097	237	294
24	6773005J1	33	637
24	6765966J1	33	606
24	6768978J1	33	631
24	g2003419	45	421
24	g1551472	61	213
24	6147606H1	71	625
24	g615579	115	462
24	g389770	122	510
24	6888770J1	153	753
24	g615989	174	503
24	4943311H1	175	458
24	4943311F6	175	595
24	6818987H1	197	267
24	1853628H1	181	421
24	1456735H1	208	332
24	5920291H1	208	267
24	7290834H1	187	505
24	6818987J1	33	250
24	6818431J1	33	570
24	g2003054	31	344
24	6770575J1	35	555
24	g1192915	25	170
24	g1978747	1	307
24	g5553287	1	315
24	6989857H1	1	436
24	6955370H1	22	540
24	g4390046	24	500
24	g4534562	24	504
25	7177245H2	1	455
25	g3015541	154	2103
25	g1864084	221	2759
25	g694473	448	790
25	g710265	448	736
25	g900615	470	914
25	g900616	469	798
25	4720263F6	580	1018

TABLE 4 (cont.)

25	4720263H1	582	820
25	g6142053	718	1125
25	g3095833	754	886
25	7213511H1	762	1242
25	g705775	879	1219
25	g1275210	960	1173
25	6551517H1	1098	1692
25	096164H1	1151	1387
25	5451192H1	1222	1451
25	1308461F6	1230	1655
25	1308461H1	1230	1360
25	385195H1	1364	1640
25	3415579H1	1387	1650
25	g1191407	1788	1959
25	4765883H1	2166	2412
25	4760585H1	2225	2489
25	1308461T6	2272	2720
25	658904H1	2278	2532
25	g2987356	2301	2759
25	g2987355	2304	2759
25	4720263T6	2360	2746
25	1308461R1	2486	2759
25	g3887078	2491	2762
25	g824280	2507	2769
26	3315579H1	1	246
26	2564790H1	4	144
26	7037134H1	17	591
26	g2214897	120	460
26	70879775V1	123	576
26	70882313V1	123	561
26	70881021V1	123	654
26	70881583V1	123	700
26	3539234F6	123	536
26	3539234H1	123	348
26	g4332214	139	571
26	5204807H1	152	395
26	7066891H1	196	711
26	70882460V1	324	844
26	6559677H1	357	941
26	70881844V1	392	965
26	70879312V1	427	993
26	7239855H1	468	1020
26	g830101	474	849
26	g889334	474	843
26	6559338H1	490	770
26	6721187H1	534	1104
26	70882570V1	535	1028
26	5780844H1	542	821
26	70882690V1	558	1104
26	5780844F6	565	1096
26	2154958H1	565	667
26	70880555V1	597	1241
26	70888508V1	603	936
26	1394886F6	630	1075
26	1394886H1	630	888
26	1392996H1	630	891
26	671307H1	655	933
26	1270677H1	663	905
26	6560774H1	677	1208
26	70885252V1	693	934
26	7289657H1	729	1231
26	g2215028	736	1137
26	6945491H1	746	1269
26	70887853V1	770	894
26	70881667V1	773	1363
26	6986634H1	816	1297
26	1374120H1	825	961
26	70882560V1	830	1440
26	70880461V1	839	1433
26	4761241H1	884	1159
26	4761249H1	885	1169
26	g901677	927	1310
26	g946847	928	1263
26	g953373	928	1130
26	70818743V1	944	1123
26	70879516V1	955	1615
26	70882124V1	977	1488
26	70881307V1	1002	1476
26	70879227V1	1036	1255
26	3803043H1	1037	1326
26	3013311H1	1056	1341
26	6883273J1	1061	1663
26	3457862H1	1084	1327
26	g316332	1120	1339
26	70880271V1	1130	1719
26	70882630V1	1138	1274
26	1391847F6	1155	1647
26	1391847H1	1155	1407
26	5292536H2	1163	1394
26	70879978V1	1205	1732
26	2453848H1	1218	1444
26	1703631H1	1230	1354
26	70879064V1	1237	1843
26	70881312V1	1275	1788
26	5385719H1	1276	1432
26	4753468H1	1281	1550
26	1966807H1	1286	1555
26	70881555V1	1332	1998
26	70818654V1	1368	1926
26	1350180H1	1376	1646
26	70879359V1	1382	1871
26	6020187H1	1410	2009

TABLE 4 (cont.)

26	70881816V1	1422	2015
26	3027682T6	1438	2026
26	1394886T6	1450	2027
26	2301449H1	1455	1541
26	70885937V1	1452	1711
26	1391847T6	1461	2030
26	3447875H2	1468	1723
26	4030281T8	1479	1804
26	70881238V1	1492	2020
26	70880651V1	1539	2110
26	4061612H1	1580	1860
26	g5863332	1584	2067
26	g5111312	1587	2067
26	2877413H1	1607	1908
26	2877413F6	1607	2002
26	g3281621	1609	2068
26	70818645V1	1622	2077
26	g4535191	1624	2068
26	g3426844	1626	2067
26	g2322267	1644	2068
26	g6196543	1654	1928
26	g3134994	1660	2074
26	g2874749	1663	2068
26	2877413T6	1681	2018
26	g830043	1717	2080
26	g946801	1740	2052
26	3539234T6	1764	2255
26	g889242	1768	2079
26	g3178069	1789	2068
26	4000739H1	1795	2068
26	g1372960	1812	4328
26	g3094856	1852	2068
26	g5528202	1869	2072
26	70887416V1	1885	2293
26	g2875209	1886	2068
26	70879855V1	1958	2305
26	70882152V1	2018	2288
26	6554433H1	2886	3287
26	g5863770	4005	4350
27	5911592T6	1	523
27	5911592H1	1	290
27	5911592T8	1	473
27	5911592F8	1	569
27	5911592T9	1	473
27	5911592F6	1	565
28	g1187505	3265	3546
28	g1128275	3293	3495
28	g1507227	3296	3546
28	g899953	3306	3566
28	g1080424	3307	3542
28	962712H1	3307	3546
28	1923976H1	3314	3512
28	g2159328	3320	3551
28	g735553	3320	3545
28	g5913481	3323	3554
28	g3896209	3322	3546
28	g795225	3331	3556
28	g2185988	2435	2887
28	4716403H1	2441	2550
28	112524H1	2441	2661
28	g6142912	2452	3005
28	4582601H1	2503	2780
28	4733207H1	2515	2810
28	g1320604	2527	3046
28	3254646H1	2529	2781
28	2273834H1	2542	2797
28	2688820H1	2567	2829
28	3449902H1	2576	2832
28	g1406097	2583	3005
28	g1406068	2588	3005
28	g2703843	2588	3002
28	g1156665	2602	2792
28	852284H1	2611	2841
28	852284R6	2613	2844
28	3477842H1	2612	2706
28	g2714143	2634	3005
28	2362491H1	2657	2912
28	g1635193	2665	2792
28	552048H1	2670	2921
28	5912223H1	2682	2748
28	g3412761	2692	3005
28	3492839H1	2695	2980
28	g1507002	2710	2916
28	5041915H1	2710	2899
28	643875H1	2715	2976
28	2531919H1	2731	2885
28	g6138438	2732	3005
28	4623249H1	2732	3002
28	2890187H1	2734	2998
28	g1670564	2741	3248
28	1850848H1	2754	3062
28	g3659213	2760	3290
28	956983H1	2762	3049
28	019839H1	2786	3082
28	3813377H1	2823	3095
28	131061H1	2831	2930
28	7054832H1	2837	3406
28	804820H1	2856	3090
28	1842462H1	2878	3146
28	4792127H1	2882	3145
28	1494563H1	2882	3121
28	1753953H1	2883	3125

TABLE 4 (cont.)

28	1755130H1	2883	3092
28	3941233H1	2902	3198
28	2116653H1	2902	3193
28	2404516H1	2914	3172
28	4524703H1	2917	3027
28	g1617791	2942	3256
28	4407776H1	2934	3211
28	5186425H1	2942	3195
28	2904404H1	2942	3200
28	3144463H1	2943	3262
28	2359103T6	2953	3498
28	4652661H1	2961	3062
28	2955930H1	2977	3261
28	3115379T6	2982	3507
28	852284T6	2987	3507
28	1661229T6	2988	3505
28	3822074H1	2994	3275
28	4229083H1	2994	3263
28	3842223H1	2994	3234
28	3607528H1	2996	3166
28	g1080514	2999	3320
28	1661229F6	3011	3447
28	1661225H1	3011	3202
28	008660H1	3047	3339
28	2321285H1	3047	3289
28	g2106118	3064	3549
28	868783H1	3065	3326
28	g5176750	3073	3550
28	g2899654	3073	3546
28	g4762266	3073	3549
28	6307419H1	3080	3547
28	g4269311	3078	3549
28	g4075892	3078	3546
28	g3740929	3094	3555
28	g3897396	3097	3546
28	612568H1	3098	3355
28	g3278888	3101	3551
28	g2899655	3101	3544
28	g3744156	3103	3546
28	g2185814	3109	3552
28	6715165H1	3111	3548
28	4864862H1	3117	3405
28	1968272R6	3132	3548
28	1968272T6	3132	3501
28	1968272H1	3132	3401
28	1492449H1	3133	3347
28	g4648047	3136	3547
28	g4438953	3138	3539
28	g2751861	3143	3349
28	g572806	3150	3528
28	g672266	3150	3466
28	g879603	3150	3402
28	g876360	3151	3531
28	g830456	3151	3412
28	321502H1	3151	3397
28	337082H1	3151	3381
28	g4891955	3153	3546
28	g5658866	3163	3547
28	3023052H1	3163	3443
28	g3884073	3170	3546
28	g5325327	3330	3546
28	g1140821	3332	3546
28	2893166T6	3341	3509
28	g2204552	3349	3551
28	g1670543	3357	3546
28	g1190688	3385	3493
28	2552971H1	3401	3550
28	5907555H1	3487	3644
28	3256027H1	3561	3626
28	3256027R6	3561	3626
28	g1959467	1	63
28	076140H1	1	230
28	3400145H1	42	272
28	7166689H1	77	373
28	5513977H1	89	336
28	4970421H1	89	348
28	g6300096	153	586
28	5335382H1	256	490
28	5335373H1	257	488
28	1437260F1	264	814
28	1437260F6	264	658
28	1437260H1	264	533
28	5373320H1	290	505
28	6485087H1	404	923
28	4181761H1	414	498
28	5026859H1	610	693
28	3230444H1	616	763
28	2134545F6	767	1341
28	2134545H1	767	1022
28	265345H1	787	970
28	1437260T6	791	1270
28	3792193H1	878	1098
28	7260531H1	921	1369
28	6986910H1	986	1376
28	4447338H1	1008	1169
28	6494154R9	1031	1550
28	4832434H1	1037	1301
28	2633783H1	1037	1287
28	g1984595	1056	1311
28	2359103R6	1060	1504
28	2359103H1	1060	1314
28	5215646H1	1093	1294

TABLE 4 (cont.)

28	425878H1	1096	1306
28	288744H1	1164	1454
28	6531566H1	1238	1809
28	7191895H2	1327	1801
28	288744F1	1349	1793
28	g6140330	1356	1781
28	g6505751	1406	1704
28	7029795H1	1414	2023
28	5641161H1	1506	1745
28	4061776T6	1508	1704
28	4061776F6	1515	1875
28	4061776H1	1516	1704
28	g2106291	1517	1824
28	g1880733	1522	1738
28	g1441510	1522	1904
28	767028H1	1524	1704
28	4177249H1	1546	1816
28	g823676	1505	1807
28	g3230537	1592	2020
28	3115379H1	1620	1700
28	g3840134	1582	1751
28	109465H1	1628	1784
28	951131H1	1599	1811
28	2431313H1	1621	1683
28	2134834H1	1679	1912
28	3811087H1	1700	1965
28	3661827H1	1726	1863
28	3729456T6	1688	1751
28	g3755762	1742	1806
28	2292441H1	1742	1982
28	2293368H1	1745	1970
28	g1939049	1757	2016
28	717351H1	1759	1999
28	g827645	1759	1975
28	5845309H1	1816	1911
28	3806331F6	1820	1915
28	6736585H1	1754	1823
28	487499H1	1809	2069
28	5914004H1	1846	2125
28	6408595H1	1852	2414
28	g1523070	1921	2355
28	g900055	1922	2243
28	5019562H1	1931	2111
28	g2103229	1933	2320
28	g2204602	1939	2229
28	2501393H1	1944	2111
28	g1281535	1964	2431
28	g735660	1994	2170
28	2813574H1	2020	2303
28	2170420H1	2030	2277
28	3718831H1	2031	2320
28	4062530H1	2048	2342
28	g1190010	2075	2225
28	4151403H1	2147	2211
28	962698R2	2147	2672
28	g6301662	2147	2523
28	3716245H1	2147	2399
28	3090607H1	2147	2385
28	962698H1	2147	2367
28	2858893H1	2147	2351
28	5586368H1	2147	2348
28	2571180H1	2147	2332
28	4333921H1	2147	2350
28	6219737H1	2147	2352
28	6400836H1	2147	2227
28	g1196242	2168	2576
28	g1190446	2168	2444
28	g1832964	2172	2494
28	675502H1	2177	2446
28	3903169H1	2207	2492
28	3245445H1	2240	2454
28	g827828	2241	2461
28	4833872H1	2258	2461
28	g1273258	2260	2749
28	4833888H1	2262	2538
28	g1799398	2268	2712
28	g1406166	2268	2643
28	g1406194	2269	2631
28	5185315H1	2285	2542
28	2082955H1	2295	2598
28	6341726H1	2316	2810
28	594752H1	2355	2602
28	g942919	2366	2583
28	7249143H1	2381	2613
28	g1921577	2394	2864
28	2896518H1	2411	2658
28	g1987258	2429	2848
28	g2161140	2435	2928
28	g3430807	3172	3546
28	6737055H1	3179	3546
28	2118476H1	3179	3436
28	5511767H1	3182	3389
28	2782179F6	3201	3588
28	2782195H1	3201	3468
28	3526177H1	3202	3479
28	g4990081	3213	3546
28	3734501H1	3227	3528
28	g3043004	3236	3546
28	g1200843	3238	3546
28	g1243436	3243	3545
28	896988R1	3244	3546
28	896988H1	3245	3472

TABLE 4 (cont.)

28	g4330537	3255	3553
28	g883772	3264	3559
29	2837088H1	1	79
29	382301H1	11	278
29	382301R6	11	248
29	381716R1	11	488
29	6853095H1	18	566
29	3296833H1	24	294
29	492559R1	36	582
29	492554H1	36	280
29	6710369H1	84	612
29	g770845	381	657
29	6710369J1	556	1057
29	6866894H1	767	1363
29	2045879F6	814	1144
29	2045879H1	814	1085
29	g677645	874	1174
29	g570913	874	1259
29	g878213	875	1218
29	3637810H1	925	1212
29	3637810F8	926	1371
29	5516287H1	958	1216
29	310657H1	1003	1205
29	054856H1	1048	1292
29	2676843H1	1123	1318
29	2865460H1	1206	1437
29	5983503F8	1245	1610
29	5983503H1	1247	1545
29	6540006H1	1281	1578
29	3903656H1	1312	1525
29	2554026H1	1346	1615
29	g1894266	1350	1824
29	7039759H1	1414	1941
29	6481201H1	1452	1566
29	6929893H1	1484	1917
29	160750H1	1667	1758
29	6201684H1	1683	2203
29	2684917H1	1733	1978
29	3898190H1	1945	2241
29	5983503T8	1966	2626
29	5952437H1	1989	2278
29	3637810T9	2048	2597
29	3151953H1	2057	2297
29	6357422H1	2085	2377
29	382301T6	2092	2657
29	2498615F6	2107	2537
29	2498615H1	2107	2341
29	492559F1	2134	2696
29	381716F1	2136	2696
29	4701147H1	2164	2436
29	g5435909	2244	2701
29	7067611H1	2285	2803
29	g2563607	2313	2696
29	1889064H1	2331	2615
29	5762206H1	2333	2712
29	2400488H1	2334	2587
29	g817549	2339	2706
29	g566965	2376	2696
29	g1894154	2387	2696
29	g869609	2428	2705
29	g4291206	2430	2805
29	g646309	2432	2696
29	7214349H1	2497	2879
29	3249908H1	2502	2799
29	672907H1	2553	2696
29	672763R6	2553	2696
29	672763H1	2553	2696
29	672696H1	2553	2696
29	672763T6	2553	2659
30	6572615H1	1	572
31	6991082H1	1	215
31	g4195018	4	167
31	g5444909	10	139
31	g5765521	10	480
31	g4736683	10	469
31	g5110384	10	474
31	g5744052	26	461
31	7181281H1	31	570
31	3801178H1	71	269
31	6606927H1	91	475
31	5725556H1	402	875
31	6459774H1	790	1082
32	g3744008	2026	2487
32	g3843455	2032	2490
32	g4334045	2035	2487
32	1295257F1	1686	2102
32	1295579H1	1686	1944
32	1295615H1	1686	1932
32	1295257H1	1686	1914
32	g1382787	1690	2060
32	3009590H1	1709	2019
32	g1327091	1710	2099
32	1496765H1	1766	2002
32	4604681H1	1772	2045
32	1596414H1	1772	1993
32	6413696H1	1785	2102
32	4534504H1	1813	2098
32	71227864V1	1847	2362
32	2210129H1	1863	2101
32	1447743H1	1866	2103
32	70861405V1	1894	2228
32	70861649V1	1895	2495

TABLE 4 (cont.)

32	6846658H1	1908	2107	32	g2657562	2083	2489	32	70793876V1	950	1625		
32	4534504T1	1907	2456	32	g5631144	2082	2483	32	71228166V1	983	1533		
32	4198839H1	1920	2101	32	70861820V1	2094	2484	32	3809253H1	1007	1304		
32	1738412T6	1927	2437	32	g4534051	2102	2483	32	1617271H1	1066	1279		
32	1737079H1	1932	2060	32	g653111	2102	2485	32	2863928H1	1081	1360		
32	1738412H1	1932	2053	32	g2741121	2113	2483	32	3234412H1	1087	1342		
32	g776871	1597	1846	32	g3900137	2112	2489	32	70861726V1	1187	1695		
32	2477944H1	1596	1816	32	g4987139	2120	2488	32	6999153H1	1207	1857		
32	4250426H1	1611	1861	32	g1327037	2121	2495	32	71228213V1	1213	1797		
32	2920084H1	1623	1883	32	g3750723	2123	2491	32	754707H1	1226	1478		
32	70862374V1	1651	2227	32	1712684T6	2129	2444	32	70860887V1	1228	1790		
32	3602331H1	1634	1931	32	5900418H1	2135	2462	32	71228275V1	1235	1732		
32	6868176H1	1636	2103	32	5900174H1	2134	2421	32	70861627V1	1248	1846		
32	4675720H1	1639	1854	32	6811079J1	1	540	32	3807022H1	1269	1442		
32	1561242F6	1658	2077	32	60205155U1	12	248	32	2950342H1	1277	1544		
32	1561242H1	1658	1879	32	6886573J1	39	560	32	2952767H1	1277	1536		
32	g1501696	1667	1973	32	6886573H1	111	596	32	71227990V1	1298	1936		
32	g760301	1677	1915	32	6811079H1	185	755	32	71228136V1	1303	1784		
32	g3278095	2137	2493	32	1453667F1	262	721	32	71227553V1	1310	1757		
32	5900945H1	2134	2423	32	1453667H1	262	526	32	70861671V1	1323	1920		
32	g6138412	2137	2496	32	1453667F6	262	546	32	70794764V1	1336	1702		
32	g4330820	2257	2483	32	70818382V1	262	390	32	3140045H1	1338	1625		
32	g1988368	2268	2493	32	3747731H1	327	524	32	70864551V1	1353	1859		
32	g3843397	2293	2490	32	973584H1	340	620	32	6210975H1	1357	1668		
32	g3920269	2298	2486	32	4043303H1	376	512	32	g653225	1358	1597		
32	4069039H1	2330	2505	32	857173H1	550	783	32	70862132V1	1378	2033		
32	g6475333	2337	2487	32	6258691H1	598	695	32	4701559H1	1384	1655		
32	312604H1	2371	2483	32	3408105H1	614	890	32	7159432H1	1388	1905		
32	313091H1	2371	2483	32	6606911H1	661	1207	32	2109285H1	1398	1660		
32	313091R6	2371	2483	32	4579377H1	669	938	32	70864775V1	1403	2064		
32	311262H1	2371	2483	32	3232119H1	686	966	32	70863822V1	1406	2038		
32	313091T6	2371	2444	32	4142126H1	717	926	32	7343876H1	1408	2057		
32	g794966	2420	2488	32	g3405461	764	1127	32	1679948H1	1413	1645		
32	5585271H1	2056	2170	32	70818359V1	915	1488	32	6210776H1	1438	1754		

TABLE 4 (cont.)

32	3866536H1	1442	1582
32	1712684F6	1443	1998
32	1712684H1	1443	1662
32	g758871	1444	1620
32	4426067H1	1466	1711
32	70795476V1	1472	1640
32	5599333H1	1493	1727
32	70797042V1	1502	1640
32	6835201H1	1538	2080
32	70863377V1	1540	1989
32	6844445H1	1560	2067
32	5155068H1	1560	1818
32	g852973	1573	1906
32	g851729	1573	1861
32	g793415	1573	1781
32	6124452H1	1584	2062
32	g788826	1597	1904
32	2130055H1	2435	2493
32	4238420H1	1936	2082
32	g2138791	1962	2385
32	4351833H1	1979	2053
32	71225822V1	1984	2102
32	71225814V1	1981	2104
32	g4390230	2003	2493
32	g4738336	2009	2484
32	g4902383	2012	2483
32	71228259V1	2018	2229
32	g4436056	2019	2491
32	71227844V1	2018	2304
32	g6037828	2021	2487
32	g3740552	2022	2489
32	g3418190	2137	2493
32	g3213525	2137	2487
32	1561242T6	2136	2435
32	g2139296	2137	2481
32	g1382788	2139	2484
32	1453667T6	2144	2442
32	g1501595	2147	2497
32	4401648H1	2175	2229
32	g760248	2190	2477
32	g3249913	2212	2489
32	g852879	2240	2477
32	g4509561	2255	2483
32	6532986H1	2257	2483
33	g779790	1220	1417
33	6117455H1	1343	1638
33	4733091H1	1405	1663
33	2614356H1	1420	1671
33	2614355H1	1420	1569
33	1340369F6	1474	1756
33	1340369H1	1474	1661
33	70920240V1	1488	2070
33	757294H1	1551	1778
33	2658667H1	1624	1866
33	2771444H1	1749	1989
33	1312886F6	1751	2202
33	1312886H1	1751	1949
33	2308711H1	1755	1965
33	3519383H1	1755	1939
33	2306567H1	1756	1936
33	1304465H1	1765	2003
33	5172484H1	1779	2028
33	4172237H1	1810	2077
33	2877775H1	1839	2116
33	869079H1	1839	2071
33	3939024H1	1856	2135
33	71273416V1	1860	2454
33	1420994H1	1918	2156
33	70917213V1	1926	2485
33	1420994F6	1937	2433
33	2661285H1	1939	2207
33	1690542H1	1958	2166
33	4044243H1	1965	2248
33	g841565	1971	2225
33	4633881H1	2015	2270
33	587465H1	2060	2372
33	756115R1	2094	2667
33	756115H1	2094	2348
33	3465750H1	2098	2249
33	71274483V1	2113	2783
33	6609076T2	2142	2819
33	71272794V1	2155	2817
33	3927045H1	2179	2474
33	3928245H1	2179	2470
33	3674253T9	2226	2768
33	2658953H1	2242	2504
33	70920349V1	2261	2805
33	4735215H1	2262	2523
33	1294470T6	2271	2833
33	2791572T6	2319	2835
33	5058201H2	2320	2433
33	1420994T6	2346	2837
33	1312886T6	2355	2836
33	1430732H1	2353	2616
33	2791668T6	2357	2837
33	2791572F6	645	894
33	6828289J1	663	1310
33	70919806V1	671	1312
33	124724H1	738	882
33	g652789	805	1068
33	2251573H1	819	1077
33	71274255V1	948	1609

TABLE 4 (cont.)

33	70920002V1	965	1599
33	70919147V1	975	1630
33	70920073V1	974	1610
33	70917224V1	1001	1557
33	g988490	1047	1351
33	71272983V1	1049	1459
33	71031330V1	1104	1535
33	4156408F6	1156	1557
33	4156408H1	1156	1423
33	71031387V1	1159	1604
33	5998189H1	1177	1292
33	71273906V1	1179	1753
33	2791668F6	1216	1550
33	2791668H1	1216	1544
33	6609076H2	1	541
33	2807474H1	7	182
33	6491123H1	19	165
33	6783159H1	27	590
33	g1727301	32	157
33	6828289H1	438	965
33	3674253H1	471	632
33	6953528H1	597	886
33	70917171V1	645	1168
33	2791572H1	646	934
33	756115F1	2364	2872
33	g5658477	2374	2795
33	g2324579	2375	2789
33	2748719H1	2415	2696
33	g4533354	2425	2876
33	g4564567	2440	2876
33	4829083H1	2441	2731
33	g5528721	2457	2877
33	g788300	2535	2872
33	g4283575	2524	2872
33	g4892982	2537	2872
33	g2410925	2550	2875
33	g652629	2559	2857
33	5316017H1	2581	2854
33	5316857H1	2585	2854
33	5318171H1	2597	2854
33	g2337727	2598	2873
33	756115T6	2617	2848
33	4735116H1	2631	2876
33	1365975R6	2632	2872
33	1365975H1	2632	2872
33	1365975T6	2633	2853
33	g1211220	2687	2875
33	2560064H1	2725	2872
33	g988325	2753	2845
34	3373528H1	609	720
34	g5754867	731	968
34	2045586H1	1036	1288
34	6799054H1	1	622
34	6452403H2	29	524
34	g1978677	101	420
34	6982612H1	143	724
34	3359232H1	147	369
34	6834663H1	387	1001
34	7001130H1	504	866
34	7318752H1	574	1174
35	1999073H1	4939	5184
35	g4330742	4944	5258
35	4934920H1	4945	5258
35	g4393289	4948	5263
35	1659543H1	4959	5214
35	g3118267	4973	5261
35	g5849381	4977	5259
35	g1218351	4988	5256
35	3130050H1	4980	5253
35	6342848H1	4981	5253
35	g866163	4979	5254
35	143138F1	4992	5258
35	g3755072	4993	5261
35	g880989	4994	5263
35	g877984	5006	5255
35	1749391T6	4740	5217
35	1344542H1	4747	5062
35	g5176036	4752	5258
35	5595877H1	4753	4917
35	6505354H1	4757	5265
35	1880971T6	4758	5218
35	g5675620	4765	5258
35	g4372792	4767	5256
35	g4281732	4769	5257
35	g5810326	4772	5259
35	g4999023	4773	5253
35	5097726H1	4779	5029
35	5685655H1	4778	5025
35	g3086706	4784	5259
35	g3752346	4790	5264
35	2183473H1	4792	5046
35	g3016110	4805	5260
35	6751216H1	4811	5148
35	5325018H1	4813	5082
35	5321404T9	4813	5124
35	5323707H1	4813	5089
35	5321503H1	4813	5077
35	g5921006	4814	5258
35	5477528H1	4813	5119
35	5482768H1	4813	5046
35	5475712H1	4813	5014
35	5323312H1	4813	5048

TABLE 4 (cont.)

35	g5511339	4816	5258	35	g847184	4909	5228	35	g434467	4329	4560		
35	g6036549	4817	5262	35	219842 3T6	4911	5218	35	1749391F6	4332	4392		
35	6337194H1	4817	4949	35	7063034H1	4916	5253	35	1749391H1	4332	4386		
35	g6399777	4829	5263	35	5485489H1	4916	5210	35	701985H1	4412	4611		
35	g6117467	4829	5264	35	1690630H1	4920	5157	35	4407419H1	4419	4685		
35	g4435700	4839	5258	35	g5888136	4922	5258	35	4708563H1	4446	4698		
35	g5636554	4842	5258	35	723564H1	4923	5070	35	6852905H1	4459	5027		
35	g3594269	4843	5258	35	723580H1	4923	5158	35	6264623H1	4490	5031		
35	g4073072	4859	5258	35	g1860289	4937	5258	35	3640801H1	4504	4758		
35	g2458074	4844	5260	35	1568070H1	4938	5172	35	2744645H1	4504	4757		
35	g4533318	4845	5258	35	2775811H1	4069	4341	35	1879458H1	4505	4778		
35	g2987667	4847	5212	35	2836761H1	4073	4337	35	7287970H1	4530	5048		
35	1924391R6	4847	5258	35	g2070265	4078	4492	35	6333393H1	4550	5092		
35	1924391T6	4847	5218	35	6812440H1	4090	4428	35	144995H1	4591	4772		
35	1924391H1	4847	5074	35	6812440J1	4090	4428	35	3147774H1	4595	4831		
35	g2555756	4854	5257	35	3151404H1	4109	4352	35	6329285H1	4599	5271		
35	g2054443	4858	5258	35	6033478H1	4116	4485	35	661058H1	4600	4880		
35	5771260H1	4874	5258	35	g2878580	4117	4402	35	1834059R6	4601	5054		
35	2246911H1	4872	5159	35	g3050962	4115	4372	35	1834059H1	4601	4873		
35	g1267895	4883	5266	35	6273920H2	4141	4414	35	6158436H1	4618	4903		
35	1339830H1	4883	5135	35	1701815H1	4143	4330	35	1622370H1	4620	4876		
35	g5590233	4886	5257	35	6426867H1	4156	4711	35	g1423847	4624	4905		
35	g4303732	4888	5259	35	6427663H1	4178	4711	35	4576478H1	4629	4893		
35	g2054335	4892	5260	35	3368975H1	4198	4330	35	600650H1	4631	4922		
35	6722884H1	4893	5253	35	g4125826	4225	4670	35	3316972H1	4638	4904		
35	g1471105	4897	5262	35	1531459H1	4225	4418	35	6954952H1	4640	5237		
35	g2963543	4895	5261	35	2966424H1	4227	4330	35	2759067H1	4643	4939		
35	g4900893	4896	5263	35	2684363H1	4237	4393	35	555514H1	4650	4902		
35	g775422	4901	5265	35	2116137H1	4273	4382	35	5334364H1	4650	4864		
35	g5362828	4902	5258	35	669344H1	4290	4560	35	5334363H1	4650	4806		
35	g5396797	4907	5264	35	2672272H1	4314	4418	35	g1367753	4648	5254		
35	3164806H1	4904	5221	35	1453860H1	4326	4539	35	3526337H1	4662	4986		
35	g5768150	4907	5251	35	1453827H1	4326	4491	35	4864025H1	4665	4953		
35	2252371H1	4909	5155	35	6179108H1	4330	4609	35	3803045H1	4668	4966		

TABLE 4 (cont.)

35	4002622H1	4679	4784
35	836008H1	4687	4806
35	2957630H1	4690	4989
35	2954183H1	4690	4974
35	6202637H1	4712	5026
35	6202437H1	4710	5128
35	2264722H1	4710	4941
35	2264938H1	4710	4910
35	g3675124	4711	5225
35	6862550H1	4721	5249
35	4941757H1	4711	5007
35	1478716H1	4711	4940
35	1476588H1	4711	4915
35	1476596H1	4711	4914
35	143138H1	4717	4918
35	145092H1	4717	4897
35	g395766	4724	5078
35	1834059T6	4728	5218
35	6393179H1	4736	5021
35	6386330H1	4737	5011
35	g866953	5008	5258
35	g867451	5014	5259
35	3865585H1	5017	5263
35	g2263181	5033	5257
35	g1741383	5041	5258
35	g3889402	5037	5258
35	g5444119	5046	5266
35	2117462H1	5071	5195
35	917065H1	5073	5258
35	g5637280	5073	5257
35	917065T1	5073	5239
35	g2464570	5078	5258
35	g2016352	5088	5258
35	5022709H1	5115	5268
35	2040433H1	5115	5221
35	4018392H1	5123	5241
35	g2079096	5140	5258
35	1453775H1	5143	5258
35	6536539H1	5171	5253
35	504486H1	5177	5246
35	g5554333	1	198
35	7030014H1	75	512
35	6984009H1	91	612
35	g2224552	197	5260
35	7092379H1	285	473
35	7193755H2	513	1006
35	6776509H1	515	1049
35	660357H1	525	791
35	661029H1	525	797
35	6990425H1	538	887
35	5623310H1	656	986
35	6939255H1	673	1165
35	6776509J1	970	1578
35	5629345H1	1070	1249
35	6348743H1	1585	1860
35	6774260J1	1597	2124
35	6765277H1	1861	2427
35	6774260H1	1904	2321
35	6516341H1	2086	2424
35	7012981H1	2178	2351
35	7075422H1	2231	2823
35	7185631H1	2343	2765
35	3101228H1	2529	2835
35	6036945H1	2608	3124
35	6637659H1	2635	3204
35	7331036H1	2646	3182
35	6637659J1	2647	3193
35	7180283H1	2692	3235
35	2708492H1	2897	2999
35	6463093H1	2925	3110
35	7091379H1	2969	3492
35	g1741484	3051	3230
35	3284115H1	3094	3353
35	1517309H1	3246	3455
35	6952950H1	3295	3883
35	3216127H1	3291	3579
35	7174368H1	3332	3903
35	3402651H1	3332	3589
35	7259765H1	3388	4023
35	6604779H1	3511	3997
35	1593761H1	3512	3747
35	7107055H1	3521	3579
35	7199042H1	3532	4116
35	6988147H1	3534	3899
35	6806336J1	3535	4013
35	6806336H1	3536	3983
35	7032229H1	3569	4118
35	3120776H1	3582	3716
35	3745702H1	3587	3892
35	3745703H1	3589	3889
35	7323378H1	3724	4337
35	7032660H1	3722	4284
35	3532688H1	3747	3964
35	6534296H1	3784	4031
35	1661311H1	3802	3897
35	2198423H1	3826	3970
35	1880971F6	3829	4311
35	1880971H1	3829	4098
35	1555666H1	3881	4099
35	1517127H1	3898	4106
35	3170592H1	3940	4237
35	6808106H1	3949	4234

TABLE 4 (cont.)

35	6808106J1	3950	4234
35	7185914H1	3966	4388
35	6943659H1	3983	4468
35	g766595	3993	4326
36	4274433H1	3948	4086
36	7289132H1	2748	3156
36	3739607H1	2770	2954
36	2149153T6	2515	3015
36	g1880151	2565	2784
36	2148724T6	2583	3030
36	g5449141	2616	3056
36	1845983T6	2617	3015
36	2658150H1	2654	2950
36	g3181486	2726	3061
36	589633R6	2736	3083
36	589633T6	2736	3029
36	g3797974	2747	3063
36	6883937H1	2092	2600
36	6979204H1	2099	2630
36	g5768436	2174	2636
36	5589055H1	2255	2525
36	5589206H1	2255	2510
36	1845983R6	2276	2760
36	1845983H1	2276	2541
36	g846523	2282	2754
36	5120292T6	2319	2628
36	5771030H1	2355	2872
36	819494H1	2363	2622
36	2149153F6	2494	2777
36	2149153H1	2494	2762
36	2593534T6	5664	6026
36	2593534F6	5671	6070
36	2593534H1	5671	5908
36	g2541279	5708	6071
36	g1751107	5760	6066
36	g778115	5856	6056
36	g2876940	6002	6062
36	3219151H1	5058	5386
36	3203918T6	5064	5609
36	3739027H1	5140	5358
36	2645933H1	5149	5412
36	g3778574	5152	5629
36	g4244154	5153	5624
36	g4311781	5164	5626
36	g4175659	5175	5634
36	3620939H1	5187	5481
36	5113889H1	5186	5447
36	2656336T6	5202	5577
36	5700054H1	5204	5442
36	5700086H1	5204	5267
36	g1751351	5217	5521
36	1679842T6	5226	5584
36	1679842F6	5233	5624
36	1679842H1	5233	5434
36	g2659077	5240	5584
36	g5813116	5263	5626
36	g2659410	5288	5628
36	g4148675	5303	5627
36	g2051261	5311	5630
36	1234495H1	5320	5628
36	2188493H1	5320	5600
36	2683448T6	5334	5590
36	g840575	5338	5626
36	7245834H1	3231	3438
36	824598R6	3289	3534
36	891226H1	3289	3534
36	824598H1	3289	3534
36	824598T1	3289	3494
36	824598T6	3289	3492
36	g2047298	3323	3838
36	g2047291	3323	3820
36	7247410H1	3362	3587
36	3203918F6	3488	3984
36	3203918H1	3489	3685
36	6172362H1	3583	3870
36	5044786H1	3736	4008
36	70046502V1	3863	4274
36	70047585V1	3863	4328
36	1304976F6	3863	4282
36	1304976H1	3863	4108
36	70047549V1	3863	4010
36	826082R1	3920	4502
36	826082H1	3920	4203
36	2308804H1	2791	3054
36	g846473	2794	3065
36	g1218558	2851	3063
36	7291393H1	2960	3486
36	6524466H1	3005	3410
36	6524566H1	3005	3543
36	1599523F6	3076	3438
36	1599523H1	3076	3277
36	g1165330	3140	3528
36	7247361H1	3207	3719
36	g1983706	3207	3474
36	3070168H1	2500	2795
36	5519150H1	4199	4369
36	2717228H1	4200	4443
36	g839478	4978	5251
36	6217349H1	4984	5467
36	2970290H1	5053	5365
36	g1982712	4550	4796
36	613186H1	4558	4795

TABLE 4 (cont.)

36	3724286H1	4560	4854
36	4365389H1	4562	4823
36	4754909H1	4583	4854
36	4354479H1	4604	4869
36	3330536H1	4650	4926
36	5581641H1	4650	4911
36	3528092H1	4659	4951
36	2750671H1	4685	4954
36	2668782H1	4691	4881
36	6372588H1	4723	4978
36	g778190	4793	5063
36	1917315H1	4825	5119
36	3621450H1	4843	5024
36	4783325H1	4844	5101
36	2656336F6	4877	5465
36	2656336H1	4877	5104
36	7336890H1	4913	5506
36	5920831H1	4918	5225
36	5096190H1	4963	5229
36	1928876H1	4970	5242
36	6217557H1	4978	5466
36	5744848H1	4239	4494
36	4176436H1	4277	4534
36	6740355H1	4458	5003
36	3487520H1	4498	4794
36	3659439H1	4514	4777
36	4274741H1	3949	4251
36	4274803H1	3949	4119
36	463357H1	4010	4201
36	4314429H1	4057	4342
36	g920351	4116	4382
36	g1149210	4133	4231
36	3766255H1	4149	4322
36	2683448F6	4167	4553
36	2683448H1	4167	4417
36	1300835T7	4174	4404
36	1307359H1	4194	4444
36	2760124H1	1934	2221
36	g858075	1936	2226
36	2760124T6	1983	2605
36	2923468H1	5441	5721
36	6838005H1	5463	5612
36	2923469T6	5476	6028
36	6838105H1	5493	5624
36	g4333756	5545	5629
36	4502184H1	5550	5622
36	5305353H1	5567	5817
36	g3647442	5625	6070
36	2733278T6	5625	6026
36	2294001H1	5633	5891
36	3993959H2	5355	5579
36	3629589H1	5367	5668
36	g2051240	5401	5630
36	1599523T6	5433	5582
36	2923469F6	5441	5868
36	2733278H1	745	977
36	g2538994	879	1084
36	7270376H1	1062	1618
36	g4242829	1103	1541
36	2780338F6	1250	1717
36	2780338H1	1250	1499
36	6244653H1	1330	1838
36	6308158H1	1771	2315
36	g2106835	1893	2201
36	2760124R6	1934	2378
36	g6330616	228	5624
36	2733278F6	745	1284
36	3994147H1	5353	5628
36	6883937J1	1	549
37	70554791V1	269	836
37	70555906V1	482	1070
37	70557145V1	488	1152
37	70328701D1	115	602
37	70557446V1	1746	2364
37	70557024V1	1777	2435
37	70326732D1	1800	2134
37	70326508D1	1800	1870
37	71304277V1	1830	2463
37	71156493V1	1852	2469
37	71303442V1	1864	2504
37	5542815H1	1873	2025
37	71157532V1	1881	2356
37	70555668V1	1893	2524
37	70555958V1	1930	2595
37	70555146V1	1931	2563
37	71303538V1	1959	2455
37	71304228V1	1958	2586
37	6496937H1	1967	2501
37	305090R6	1971	2342
37	305090H1	1970	2306
37	4598818H1	1996	2251
37	6349213H2	2054	2378
37	70556404V1	1493	2023
37	3696047F6	1521	2066
37	3696047H1	1522	1818
37	71158742V1	1536	2128
37	71156538V1	1542	2034
37	70327564D1	1550	2005
37	4670450H1	1563	1762
37	71157870V1	1598	2195
37	70556820V1	1615	2235
37	6416418H1	1667	1887

TABLE 4 (cont.)

37	6389818H1	1667	1987
37	4518860H1	1672	1933
37	70554892V1	1703	2343
37	70554965V1	1703	2332
37	6830659J1	1705	2343
37	3279857H1	1719	1993
37	71304118V1	1741	2354
37	71158362V1	1743	2480
37	71155779V1	2409	2987
37	4172634F6	2447	3014
37	4172634H1	2447	2722
37	4438947H1	2448	2716
37	71156387V1	2457	2883
37	71303533V1	2512	2939
37	7353820H1	2529	2887
37	4539057H1	2561	2815
37	2328218H1	2633	2899
37	71304436V1	2666	3213
37	71157628V1	2710	3265
37	5106567H1	2713	2961
37	4599088H1	2761	3020
37	1501621F6	2190	2690
37	1501621H1	2190	2378
37	70557357V1	2284	2914
37	71157279V1	2290	2770
37	6116935H1	2291	2555
37	70325710D1	2321	2741
37	70325612D1	2363	2756
37	70328746D1	2363	2721
37	71156954V1	2388	2865
37	761848H1	2387	2597
37	2528759H1	2396	2656
37	70555774V1	2404	3076
37	3222459H1	2408	2765
37	70555710V1	602	1210
37	70554866V1	605	1225
37	70327790D1	614	1116
37	70325412D1	620	997
37	70326955D1	620	1007
37	6828695H1	703	1285
37	2868052H1	708	843
37	70555300V1	723	1261
37	1582746H1	3153	3386
37	g5848554	3164	3419
37	2770719T6	3195	3431
37	6416515H1	3258	3419
37	g4739984	3348	3419
37	6785591H1	12	523
37	2925464F6	16	568
37	4179240H1	17	287
37	2925464H1	16	274
37	4179553F8	21	514
37	4179553H1	21	247
37	4874914H1	4	263
37	4179741H1	4	294
37	6075277H1	2826	3033
37	1426361F6	2857	3303
37	1426357H1	2857	3060
37	71131546V1	2866	3169
37	5536040H1	2910	3142
37	1501621T6	2953	3435
37	71158019V1	2958	3419
37	4050931H1	2977	3284
37	70326238D1	2988	3419
37	4179553T9	2999	3343
37	71156430V1	3001	3419
37	g4665411	3004	3419
37	4172634T6	3023	3429
37	g2099982	3028	3419
37	2770719H1	3054	3325
37	2770719F6	3054	3249
37	g2077519	3061	3419
37	g2099950	3063	3288
37	g5664324	3092	3419
37	g5452554	3115	3474
37	71158855V1	1155	1627
37	5811393H1	1155	1458
37	71157014V1	1155	1753
37	g5850365	1172	1534
37	g5865429	1177	1479
37	70446257V1	1237	1854
37	70446298V1	1236	1858
37	70326574D1	1292	1722
37	70555309V1	1308	1895
37	70555528V1	1315	1998
37	70556256V1	1368	2053
37	70556149V1	1371	1998
37	70555054V1	1382	1948
37	70555206V1	1385	1982
37	4441126H1	1384	1659
37	70557288V1	1422	2021
37	70560338V1	1426	2013
37	70326191D1	1440	1766
37	70327556D1	1458	2005
37	3699373H1	25	340
37	70327386D1	26	382
37	6784564H2	35	536
37	6786847H2	39	668
37	70554782V1	730	1378
37	70555359V1	732	1309
37	6830659H1	734	1265
37	70555879V1	743	1324

TABLE 4 (cont.)

37	70556961V1	761	1427
37	70557092V1	784	1383
37	70554523V1	792	1538
37	70557219V1	804	1427
37	70555075V1	854	1389
37	70555282V1	856	1303
37	70554784V1	862	1429
37	6785373H1	889	1448
37	70556389V1	938	1426
37	70556118V1	963	1544
37	70557489V1	1005	1631
37	70554717V1	1009	1418
37	6784929H1	1068	1464
37	6828695J1	1071	1726
37	70556000V1	1081	1742
37	6934607H1	1085	1599
37	70449057V1	1109	1224
37	71303301V1	1146	1592
37	5811393F6	1155	1729
37	71156205V1	1155	1718
37	71156521V1	1155	1693
37	70554574V1	568	1182
37	70556236V1	564	1260
37	70554808V1	577	1186
37	6788638H1	13	474
37	6787884H1	1	326
37	71303881V1	1465	2036
37	6788583H1	1	581
37	6788770H1	510	1086
37	70554811V1	2066	2662
37	4515767H1	2069	2207
37	71303748V1	2138	2612
37	70328165D1	2151	2705
37	70326303D1	2151	2673
37	70326287D1	2151	2447
37	71155657V1	2163	2702
37	4179741T9	2811	3358
37	70556579V1	2797	3121
37	71303602V1	2803	3455
37	g2051100	2822	3123
38	60100196D1	1959	2231
38	1859554H1	2167	2443
38	1859570H1	2167	2444
38	3361850H1	2214	2460
38	5272051H1	2369	2567
38	5272051F9	2369	2887
38	5272051F8	2369	2912
38	5090972F6	2471	2993
38	5090972H1	2471	2747
38	4274991F6	2519	2898
38	4274991H1	2519	2780
38	2185660H1	2581	2841
38	5090972R6	2805	3071
38	g5802614	1	3437
38	60100191D1	1682	2005
38	g1373056	1770	2132
38	6489031H1	1908	2435
38	5272051T9	2893	3324
38	4274991T6	2954	3393
38	g4196744	2957	3437
38	60100196B1	2968	3406
38	60100198B1	3119	3474
38	60100190B1	3184	3401
38	g3418913	3219	3438
38	60100191B1	3333	3472
38	196837H1	3382	3511
39	6775050J1	717	1394
39	6775050H1	925	1555
39	7361157H1	1029	1613
39	579137H1	1293	1511
39	g6197626	1359	1828
39	7156184J2	747	1335
39	7277468H1	854	1192
39	g2986601	375	462
39	5844017H1	418	618
39	7324537H1	307	843
39	g1277998	1	466
39	804517H1	25	265
39	4918488H1	31	303
39	7156184H2	35	641
39	1703886F6	35	435
39	1703886H1	35	245
39	3809668H1	45	350
39	g5152120	74	458
39	4550249H1	1	264
39	g6142263	81	462
39	g2254363	214	462
39	1703886T6	232	484
39	2656212F6	290	462
40	5314759H1	182	438
40	6222064U1	497	1056
40	g3003145	668	944
40	3818881F6	1	468
40	70536625V1	1	563
40	3818881H1	1	280
40	3345551H1	83	362
40	5988985F9	102	643
40	5988985H1	102	378
40	6267489H1	104	741
40	4072614H1	112	399
40	7167692H1	120	649
41	g1545026	2331	2704

TABLE 4 (cont.)

41	g1062645	2331	2693
41	g1064773	2331	2676
41	g1482703	2331	2498
41	6549638H1	2430	3013
41	70300848D1	2452	2708
41	70300835D1	2479	2708
41	415443H1	2572	2798
41	419855H1	2572	2791
41	416163H1	2572	2762
41	1739793H1	3085	3321
41	1739793T6	3100	3767
41	4422806H1	3205	3454
41	415443F1	3205	3806
41	70300638D1	3222	3594
41	70300351D1	3251	3666
41	1595527T6	3300	3770
41	1595527H1	3307	3511
41	415986F1	3324	3806
41	4879243H1	3381	3654
41	g6139643	3394	3806
41	g1482608	3395	3806
41	2287181H1	3404	3604
41	2287181R6	3404	3572
41	g1162076	3447	3742
41	g1527588	3504	3806
41	g1481970	3516	3806
41	5779072H1	3534	3787
41	70300150D1	3556	3802
41	g1062646	3606	3790
41	g1064735	3701	3781
41	g4112497	3100	3288
41	684750H1	3105	3340
41	2402302H1	3037	3261
41	1739793R6	3085	3458
41	2497235H1	1745	2055
41	7190840H1	2160	2660
41	3285638H1	2171	2415
41	3285638F6	2171	2570
41	70300497D1	1250	1823
41	3348848H1	1522	1695
41	60133508V1	1520	1825
41	60131087B1	2209	2545
41	70300222D1	2312	2702
41	g1482020	2331	2775
41	2897538H1	1	259
41	g5457042	169	2567
41	3901248T9	378	1003
41	3899909T8	440	979
41	70516717D1	1091	1389
41	70300884D1	1130	1406
41	415991H1	2572	2642
41	415443R1	2572	3083
41	6362320H1	2606	2807
41	2783446H2	2623	2867
41	4442155H1	2651	2857
41	1849376H1	2685	2967
41	3285638T6	2784	3315
41	70300827D1	2787	3376
41	g3447015	2817	3261
41	2879330H1	2863	3165
41	g4110893	2891	3343
41	g6037968	2932	3343
41	g3693629	2952	3343
41	4113890H1	2979	3246
41	70300837D1	1134	1556
41	70300823D1	1230	1552
41	60211594U1	1243	1746
42	70866933V1	5034	5705
42	5926529H1	5081	5401
42	g1751265	5091	5420
42	4767333H1	5123	5429
42	70812418V1	5132	5800
42	5833936H1	5148	5428
42	g3016077	5152	5415
42	g4149219	5242	5421
42	70814699V1	5281	5854
42	70868813V1	5288	5908
42	1373555H1	5301	5546
42	g4307618	5322	5811
42	70867023V1	5332	5966
42	70869633V1	5404	6021
42	g2409915	5411	5811
42	1433020H1	5460	5705
42	70867216V1	5558	6222
42	1267718H1	4756	5019
42	g318200	4774	5165
42	1464866H1	3731	3992
42	70870570V1	3750	4457
42	71230331V1	3765	4290
42	71222361V1	3780	3934
42	71190090V1	3794	4487
42	70837174V1	3808	4000
42	71216238V1	3533	4246
42	71189613V1	3573	4128
42	4147558H1	4630	4860
42	71191702V1	4644	5197
42	3769383H1	4647	4965
42	71131533V1	4665	5137
42	70816797V1	4715	5387
42	71188635V1	4724	5165
42	7051349H1	3739	4208
42	71189574V1	4075	4700

TABLE 4 (cont.)

42	g612859	4079	4410
42	71189238V1	4084	4700
42	g570718	4094	4400
42	71188405V1	4111	4753
42	g2805702	4165	4597
42	g3694501	4167	4598
42	71189379V1	4173	4848
42	g6144708	4176	4598
42	g2323168	4177	4598
42	g819401	4184	4610
42	70868193V1	4190	4727
42	g766671	4190	4568
42	g1516806	4197	4665
42	g1525425	4197	4612
42	g830693	4218	4610
42	71188787V1	4238	4612
42	4785755H1	4253	4533
42	70866811V1	4297	4860
42	g1614228	4303	4568
42	g3229742	467	888
42	g5457022	725	3257
42	g5456921	725	6222
42	g4683485	1334	1781
42	g5765573	1334	1759
42	g3075910	1387	1688
42	7190218H2	2401	2913
42	71229788V1	2815	3413
42	5014904F6	2815	3221
42	5014904H1	2815	3090
42	71229920V1	2973	3658
42	71228807V1	3182	3779
42	6884462H1	3181	3686
42	70868094V1	3257	3948
42	70869027V1	3256	3892
42	g823731	3267	3515
42	7044511H1	3272	3873
42	5919091H1	3287	3555
42	71188683V1	3333	3897
42	71191815V1	3333	3961
42	71191533V1	3333	3856
42	71191734V1	3333	3854
42	1600316F6	3333	3729
42	1600316H1	3333	3435
42	70867333V1	3341	3911
42	g839823	3355	3689
42	g824451	3355	3650
42	71190867V1	3363	3882
42	70870265V1	3396	4051
42	2013807H1	3391	3501
42	70866888V1	3809	4505
42	3673862H1	5564	5859
42	2499983T6	5584	6176
42	g815044	4346	4627
42	70869526V1	4360	4860
42	2499983H1	4367	4635
42	70867729V1	4376	5138
42	5386383H1	4385	4647
42	70868265V1	4388	5095
42	6274578H1	4416	4860
42	71190615V1	4434	5068
42	70866931V1	4477	5082
42	71189990V1	4513	5134
42	71190387V1	4513	5133
42	g672203	4544	4860
42	1269521F6	4577	5030
42	1269521H1	4577	4812
42	71230051V1	4584	5171
42	g670126	4590	4860
42	71188785V1	4601	5195
42	1817860T6	4773	5373
42	71230388V1	3476	4062
42	6337414H1	4795	5436
42	71188365V1	4864	5408
42	71129972V1	4882	5273
42	g3887571	4884	5422
42	7052610H1	3740	3875
42	71230123V1	4787	5357
42	1600316T6	4788	5379
42	g2224630	1	6155
42	g2142053	464	854
42	g3842828	466	883
42	1311611F6	4886	5420
42	1311611T6	4886	5378
42	g575078	4886	5176
42	1311611H1	4886	5148
42	71188609V1	4890	5438
42	71229950V1	4890	5346
42	2293604H1	4890	5151
42	621828H1	4890	5148
42	2626661H1	4890	5070
42	1269521T6	4892	5380
42	6327560H1	4893	5348
42	g2539162	4894	5429
42	g4852194	4905	5421
42	g2932593	4922	5424
42	g3148673	4928	5422
42	7098720H1	4931	5587
42	g5707120	4951	5413
42	3975608H1	4953	5272
42	3975908H1	4954	5274
42	70814653V1	4965	5676
42	g4971769	4971	5424

TABLE 4 (cont.)

42	71188351V1	3626	4086
42	70838919V1	3631	4136
42	71188254V1	3638	4239
42	71189595V1	3651	3907
42	70870573V1	3655	4351
42	70868067V1	3679	4334
42	70867164V1	3682	4354
42	71230406V1	3685	4227
42	70869964V1	3682	4340
42	1817860F6	3725	4287
42	1817860H1	3725	4029
42	7050051H1	3739	4283
42	70816308V1	4604	5347
42	70813062V1	4615	5238
42	7103719H1	4627	5050
42	g883091	4613	5038
42	1963922R6	4615	5216
42	70825247V1	4615	5083
42	70815988V1	4615	5030
42	70649447V1	4615	5280
42	70814603V1	4615	5185
42	70812386V1	4615	5163
42	70813116V1	4615	5137
42	70812591V1	4615	5112
42	1963922H1	4615	4860
42	70817149V1	4615	5238
42	71190973V1	3394	4015
42	70866857V1	3421	4053
42	70869712V1	3422	4110
42	71190024V1	3462	4134
42	71222510V1	3809	4002
42	71229550V1	3828	4582
42	7317184H2	3840	4515
42	71191575V1	3866	4388
42	71190157V1	3909	4588
42	71230422V1	3911	4602
42	70868868V1	3926	4435
42	71191209V1	3938	4502
42	71229173V1	3944	4466
42	71191826V1	3939	4349
42	71188071V1	3982	4494
42	71222526V1	3993	4351
42	70868437V1	4000	4529
42	70867683V1	4003	4658
42	71190956V1	4017	4607
42	70867083V1	4019	4527
42	70869984V1	4019	4488
42	g775853	4047	4392
42	71189002V1	4049	4491
42	70870114V1	4057	4751
42	1963922T6	5617	6180
42	745052H1	5643	5869
42	3333795T6	5681	6181
42	4421884H1	5703	5956
42	g4989315	5743	6225
42	g3446159	5744	6227
42	g5853840	5747	6219
42	2280040T6	5748	6175
42	g4264936	5749	6222
42	g5590548	5767	6219
42	2280040R6	5769	6222
42	2280040H1	5769	6044
42	g4114692	5775	6229
42	2157793H1	5776	6020
42	g4269881	5783	6222
42	g314938	5790	6222
42	5014904T6	5789	6175
42	g1516807	5846	6222
42	71190271V1	3599	4339
42	5515021R7	3622	4216
42	71229150V1	3622	4275
42	70867419V1	3623	4261
42	g671390	5960	6219
42	g820781	5971	6244
42	g668623	6031	6222
42	71221653V1	6103	6222
42	g882914	6021	6129
42	71188120V1	4750	4951
42	1267718F1	4756	5198
42	71190911V1	4733	5379
42	71188586V1	4756	5397
42	70869357V1	4982	5696
42	g3756453	4981	5424
42	4776237H1	4985	5261
42	71190506V1	5033	5514
42	6608393T1	5498	6138
42	5907377H1	5524	5800
42	70870592V1	5528	6173
42	70813957V1	5544	6036
42	3333795F6	5552	6027
42	3333795H1	5552	5840
42	71188885V1	4599	5206
42	g1525426	5842	6222
42	g882983	5853	6245
42	g797506	5865	6230
42	g587184	5880	6222
42	70870719V1	5924	6239
42	g814957	5894	6223
42	g822523	5964	6230
42	g612999	4719	5074
43	g2034169	2102	2394
43	5540505T7	2291	2870

TABLE 4 (cont.)

43	6377332H1	2417	2702
43	4947810H1	2612	2733
43	g5006247	1	2762
43	5540505F6	953	1415
43	5540505H1	953	1146
43	g2875734	2835	2940
43	g3735348	2634	2945
43	5118201T6	2631	2910
43	2749265F6	2448	2923
43	2749265H1	2448	2714
43	2749265T6	2551	2897
43	537065H1	2429	2663
44	1452312F1	3288	3835
44	70007188D1	3260	3637
44	g898311	3282	3460
44	1452312F6	3288	3736
44	1452312H1	3288	3560
44	2599007H1	3312	3589
44	6325947H1	3442	3749
44	840648H1	3415	3672
44	70012088D1	3420	3797
44	5852153H1	3426	3701
44	70604010V1	1419	2043
44	6952285H1	1480	2049
44	4458494F6	1493	1942
44	70608095V1	1492	1936
44	4458494H1	1494	1730
44	7255931H2	1571	1752
44	6909665J1	1608	2154
44	6969377U1	1616	2026
44	2272356R6	1622	1941
44	2272356H1	1622	1890
44	70608114V1	1801	1904
44	6553230H1	1811	2165
44	6559394H1	1811	2428
44	3382113H1	1881	2090
44	70606021V1	1880	2259
44	70879980V1	2089	2579
44	2661806F6	2089	2531
44	2661806H1	2089	2361
44	70879113V1	2089	2545
44	g6476309	2149	2506
44	2627073H1	2160	2391
44	2627315H1	2160	2389
44	3901711H1	2247	2491
44	70887530V1	2263	2344
44	6969302U1	2280	2623
44	70881572V1	2297	2821
44	5763849H1	2351	2873
44	7256511H1	2398	2905
44	70882796V1	2405	3030
44	70886211V1	2434	2594
44	70882791V1	2477	2906
44	70882271V1	2478	2974
44	70881365V1	2478	2973
44	70003939D1	2481	2947
44	70012299D1	2481	2829
44	70004016D1	2481	3025
44	3572311F6	2487	3077
44	3572311H1	2487	2699
44	70005627D1	2487	2687
44	70010847D1	2517	2952
44	7336064H1	2527	2982
44	70880257V1	2544	3145
44	70011933D1	2553	3044
44	2272356T6	2566	3001
44	70888761V1	2568	2873
44	3011048H1	3342	3641
44	4562117H1	3350	3613
44	4563263H1	3352	3636
44	70603379V1	1131	1723
44	70603933V1	1153	1782
44	70607414V1	1277	1412
44	70607363V1	1042	1396
44	2414751H1	3218	3489
44	389997H1	3676	3915
44	6357624H1	3682	3922
44	g3961665	3684	3920
44	g6477150	3686	3925
44	1689958F6	3693	3923
44	1689958H1	3693	3907
44	1689958T6	3698	3880
44	1702166T6	3718	3866
44	3572311T6	3740	3872
44	g4649451	3791	3915
44	4099042H2	3816	3927
44	4099042F8	3816	4438
44	1243554H1	3816	3923
44	g4325490	3834	3915
44	2968601H1	3954	4247
44	g5810032	3494	3926
44	7255223H1	3518	3915
44	g2237335	3527	3920
44	2878117H1	3530	3815
44	g1400734	3536	3915
44	5104505H1	3540	3772
44	g4081742	3542	3923
44	1452312T6	3546	3876
44	g898312	3565	3918
44	6499719H1	3564	3909
44	g4081564	3565	3923
44	g2335900	3599	3920

TABLE 4 (cont.)

44	g6451467	3602	3915
44	g1521304	3605	3931
44	g4534027	3606	3923
44	5790863H1	3609	3903
44	5789451H1	3609	3898
44	5787849H1	3609	3915
44	g5528373	3621	3920
44	g1516463	3624	3931
44	g5912966	3660	3920
44	344685H1	3673	3922
44	2623608H1	3367	3604
44	840648R1	3415	3915
44	4333836H1	3415	3703
44	70881547V1	3400	3921
44	70886619V1	3404	3634
44	2414749F6	3218	3747
44	70605048V1	1033	1331
44	7267489H1	1034	1578
44	6346421H1	3442	3736
44	6317150H1	3442	3746
44	4897563H1	3129	3422
44	5379052H1	3137	3362
44	3406784H1	3145	3410
44	70008878D1	3156	3637
44	70608052V1	1080	1187
44	g3888759	1108	1488
44	2857322H1	2904	3183
44	70881851V1	2904	3275
44	792748R1	2910	3533
44	792748H1	2909	3154
44	793130H1	2910	3134
44	7159471H1	2922	3506
44	70880131V1	2923	3534
44	1541872H1	2940	3161
44	684595H1	2941	3207
44	70886274V1	2982	3197
44	70886318V1	2982	3196
44	6722223H1	3013	3202
44	2806050H1	3019	3347
44	1702166F6	3044	3568
44	1702166H1	3044	3271
44	4980587H1	3057	3327
44	6909665H1	3076	3619
44	4372755H1	3078	3384
44	6074761H1	3079	3396
44	685902H1	2605	2829
44	70880726V1	2616	3181
44	2615527H1	2623	2881
44	70879436V1	2671	3129
44	70882269V1	2673	3180
44	70887568V1	2676	2818
44	70882659V1	2688	3179
44	1438876F1	2686	3071
44	1438880H1	2686	2970
44	1438876H1	2686	2968
44	2258046H1	2717	2963
44	70003496D1	2721	3284
44	70011398D1	2733	3192
44	70882502V1	2739	3418
44	70879669V1	2748	3253
44	70006402D1	2745	3309
44	70004115D1	2745	3108
44	70011055D1	2745	3198
44	70882244V1	2768	3039
44	70007592D1	2769	2981
44	6479471H1	2787	3356
44	7054594H1	2797	3403
44	70879623V1	2807	3487
44	5274874H1	2829	3072
44	70007727D1	2843	3340
44	70010542D1	2843	3307
44	70010162D1	2843	3246
44	70005864D1	2843	3198
44	70002001D1	2843	3074
44	70002333D1	2844	3415
44	70011761D1	2844	3198
44	70001785D1	2849	3344
44	70007867D1	2874	3336
44	70006872D1	2875	3344
44	70004362D1	2885	3284
44	70604116V1	1123	1734
44	2658395H1	3490	3738
44	70879732V1	3478	3911
44	g3429071	3484	3920
44	6317128H1	3442	3575
44	70879089V1	3455	3925
44	2661806T6	3469	3883
44	700495H1	3477	3740
44	70608699V1	853	1342
44	70653541V1	904	1439
44	70607650V1	918	1337
44	6938224H1	924	1338
44	70608866V1	964	1616
44	3776430H1	3217	3522
44	709518H1	3215	3449
44	70888779V1	3218	3398
44	872814H1	3082	3286
44	5438843H1	3097	3403
44	70003362D1	3164	3424
44	70004958D1	3165	3415
44	2527855H1	3178	3528
44	g1521303	3198	3655

TABLE 4 (cont.)

44	g1517127	3198	3698	45	1524230H1	43	257			
44	2414483H1	3218	3454	45	3384786H1	92	329			
44	70010299D1	3248	3632	45	6055559H1	174	688			
44	70005831D1	3338	3877	45	6055841H1	174	688			
44	70003405D1	3101	3415	45	4509676H1	259	437			
44	70007838D1	3099	3382	45	3081417H1	405	589			
44	4880465H1	3100	3351	45	2952165H1	422	670			
44	70012577D1	3107	3637	45	70874349V1	542	987			
44	1320150H1	3127	3364							
44	70008556D1	3132	3440							
44	4181419H1	1	167							
44	6779195J1	66	705							
44	113399R6	430	794							
44	4507995F6	435	610							
44	4507995H1	436	607							
44	6831490H1	443	635							
44	6831490J1	443	635							
44	70604944V1	690	1146							
44	70607511V1	785	1414							
44	6454789H1	1287	1795							
44	70603538V1	1322	1922							
44	684735H1	1352	1601							
44	70607606V1	1355	1770							
44	70603837V1	1402	1982							
44	70006129D1	3099	3637							
45	3386984H1	1	235							
45	3087717H1	1	207							
45	4832592H1	11	232							
45	3750644H1	15	214							
45	3350574H1	18	296							
45	3150464H1	24	307							
45	3381160H1	29	281							
45	3092918H1	38	363							
45	3092958H1	38	329							

TABLE 5

SEQ ID NO:	Template ID	Tissue Distribution
1	LG:977683.1:2000FEB18	Nervous System - 21%, Skin - 19%, Embryonic Structures - 11%
2	LG:893050.1:2000FEB18	Digestive System - 40%, Hemic and Immune System - 40%, Nervous System - 20%
3	LG:980153.1:2000FEB18	Nervous System - 16%, Urinary Tract - 12%, Skin - 12%
4	LG:350398.1:2000FEB18	Digestive System - 50%, Hemic and Immune System - 50%
5	LG:475551.1:2000FEB18	Skin - 35%, Hemic and Immune System - 19%, Digestive System - 11%
6	LG:481407.2:2000FEB18	widely distributed
7	LI:443580.1:2000FEB01	Unclassified/Mixed - 60%, Connective Tissue - 17%, Endocrine System - 13%
8	LI:803015.1:2000FEB01	Urinary Tract - 63%, Respiratory System - 38%
9	LG:027410.3:2000MAY19	Respiratory System - 100%
10	LG:171377.1:2000MAY19	Unclassified/Mixed - 74%, Female Genitalia - 13%, Cardiovascular System - 10%
11	LG:352559.1:2000MAY19	Unclassified/Mixed - 71%, Digestive System - 29%
12	LG:247384.1:2000MAY19	Stomatognathic System - 39%, Musculoskeletal System - 28%, Cardiovascular System - 19%
13	LG:403872.1:2000MAY19	Nervous System - 40%, Embryonic Structures - 23%, Urinary Tract - 14%
14	LG:1135213.1:2000MAY19	Embryonic Structures - 24%, Cardiovascular System - 20%, Unclassified/Mixed - 13%
15	LG:474284.2:2000MAY19	Unclassified/Mixed - 14%
16	LG:342147.1:2000MAY19	Pancreas - 21%, Male Genitalia - 19%, Female Genitalia - 17%, Urinary Tract - 17%
17	LG:1097300.1:2000MAY19	Endocrine System - 25%, Skin - 18%, Unclassified/Mixed - 13%
18	LG:444850.9:2000MAY19	Digestive System - 28%, Connective Tissue - 20%, Exocrine Glands - 10%
19	LG:402231.6:2000MAY19	Endocrine System - 23%, Hemic and Immune System - 23%, Digestive System - 18%
20	LG:1076157.1:2000MAY19	Embryonic Structures - 50%, Endocrine System - 28%, Respiratory System - 17%
21	LG:1083142.1:2000MAY19	Germ Cells - 84%
22	LG:1083264.1:2000MAY19	Liver - 52%, Connective Tissue - 33%
23	LG:350793.2:2000MAY19	Sense Organs - 25%, Connective Tissue - 14%
24	LG:408751.3:2000MAY19	Nervous System - 39%, Sense Organs - 39%
25	LI:336120.1:2000MAY01	Nervous System - 24%, Respiratory System - 22%, Endocrine System - 18%
26	LI:234104.2:2000MAY01	Female Genitalia - 21%, Unclassified/Mixed - 17%, Nervous System - 12%
27	LI:450887.1:2000MAY01	Nervous System - 100%
28	LI:119992.3:2000MAY01	Embryonic Structures - 10%
29	LI:197241.2:2000MAY01	Connective Tissue - 26%, Endocrine System - 12%
30	LI:406860.20:2000MAY01	Digestive System - 100%
31	LI:142384.1:2000MAY01	Connective Tissue - 44%, Germ Cells - 34%
32	LI:895427.1:2000MAY01	Cardiovascular System - 20%, Urinary Tract - 14%, Skin - 13%
33	LI:757439.1:2000MAY01	Digestive System - 18%, Embryonic Structures - 13%, Sense Organs - 12%

34	LI:1144066.1:2000MAY01	Cardiovascular System - 59%, Exocrine Glands - 25%
35	LI:243660.4:2000MAY01	Pancreas - 63%
36	LI:334386.1:2000MAY01	Exocrine Glands - 17%, Male Genitalia - 16%, Musculoskeletal System - 13%
37	LI:347572.1:2000MAY01	Digestive System - 30%, Digestive System - 23%, Respiratory System - 17%
38	LI:817314.1:2000MAY01	Unclassified/Mixed - 55%, Male Genitalia - 26%, Female Genitalia - 11%
39	LI:000290.1:2000MAY01	Female Genitalia - 54% Urinary Tract - 50%, Musculoskeletal System - 27%, Hemic and Immune System - 23%
40	LI:023518.3:2000MAY01	Sense Organs - 72%
41	LI:1084246.1:2000MAY01	Musculoskeletal System - 19%, Germ Cells - 18%, Nervous System - 14%
42	LI:1165828.1:2000MAY01	Connective Tissue - 29%, Respiratory System - 21%, Hemic and Immune System - 18%
43	LI:007302.1:2000MAY01	Skin - 30%, Female Genitalia - 11%
44	LI:236386.4:2000MAY01	Exocrine Glands - 20%, Nervous System - 16%, Endocrine System - 13%
45	LI:252904.5:2000MAY01	

TABLE 6

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability score	Annotation
46	3	263	27	815	g10764778	1e-131	phosphoinositol 3-phosphate-binding protein-2 [Homo sapiens]
					g10045840	1e-58	TPC2 [unidentified]
					g4589582	2e-28	KIAA0969 protein [Homo sapiens]
47	1	217	10	660	g6634025	1e-81	KIAA0379 protein [Homo sapiens]
					g6453538	6e-77	hypothetical protein [Homo sapiens]
					g4803678	7e-29	ankyrin (brank-2) [Homo sapiens]
48	1	716	613	2760	g7243215	0.0	KIAA1417 protein [Homo sapiens]
					g7263990	0.0	dJ93K22.1 (novel protein (contains DKFZP564B116)) [Homo sapiens]
					g7302944	5e-57	CG8060 gene product [Drosophila melanogaster]
49	3	107	60	380			
50	3	645	3	1937	g4826478	0.0	dJ37E16.2 (SH3-domain binding protein 1) [Homo sapiens]
					g861029	0.0	SH3 domain binding protein [Mus musculus]
					g7018521	0.0	hypothetical protein [Homo sapiens]
51	3	177	93	623	g6119546	1e-45	hypothetical protein; 114721-113936 [Arabidopsis thaliana]
					g6522593	3e-10	putative RNA binding protein [Arabidopsis thaliana]
					g950424	4e-10	splicing factor, arginine/serine-rich 7 [Homo sapiens]
52	1	217	79	729	g4589566	3e-34	KIAA0961 protein [Homo sapiens]
					g3970712	3e-26	zinc finger protein 10 [Homo sapiens]
					g7630121	8e-25	zinc finger protein 92 [Mus musculus]
53	3	151	3	455	g5262560	2e-35	hypothetical protein [Homo sapiens]
					g10434856	1e-29	unnamed protein product [Homo sapiens]
					g930123	9e-27	zinc finger protein (583 AA) [Homo sapiens]
54	3	193	3	581	g10438267	1e-74	unnamed protein product [Homo sapiens]
					g7290756	8e-16	CG4532 gene product [Drosophila melanogaster]
					g5705877	8e-10	POD-1 [Caenorhabditis elegans]
55	3	282	3	848	g3077703	1e-111	mitsugumin29 [Oryctolagus cuniculus]
					g3461888	1e-108	mitsugumin29 [Mus musculus]
					g3761107	1e-108	mitsugumin29 [Mus musculus]

TABLE 6 (cont.)

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability score	Annotation
56	2	211	2	634	g7243243	2e-44	KIAA1431 protein [Homo sapiens]
					g4567179	2e-43	BC37295_1 [Homo sapiens]
					g3445181	1e-41	R31665_2 [Homo sapiens]
57	2	366	83	1180	g9945010	1e-120	RING-finger protein MURF [Mus musculus]
					g9929937	5e-92	hypothetical protein [Macaca fascicularis]
					g10439844	1e-36	unnamed protein product [Homo sapiens]
58	3	326	354	1331	g7020303	0.0	unnamed protein product [Homo sapiens]
					g10434892	3e-79	unnamed protein product [Homo sapiens]
					g6683707	2e-31	KIAA0455 protein [Homo sapiens]
59	1	156	70	537	g6692607	2e-69	MGA protein [Mus musculus]
					g5931585	9e-47	T-box family member; T-box domain [Cynops pyrrhogaster]
					g4049463	3e-16	transcription factor TBX6 [Homo sapiens]
60	2	262	239	1024	g1488047	7e-12	RING finger protein [Xenopus laevis]
					g3916727	1e-11	estrogen-responsive B box protein [Homo sapiens]
					g401763	1e-11	ataxia-telangiectasia group D-associated protein [Homo sapiens]
61	3	132	138	533			
62	2	167	2	502	g2078531	2e-71	Mlark [Mus musculus]
					g2078529	2e-70	Hlark [Homo sapiens]
					g1149523	8e-57	Neosin [Mus musculus]
63	1	570	160	1869	g183002	0.0	guanylate binding protein isoform I [Homo sapiens]
					g829177	0.0	guanylate binding protein isoform II [Homo sapiens]
					g7023332	0.0	unnamed protein product [Homo sapiens]
64	3	168	3	506	g7020737	2e-89	unnamed protein product [Homo sapiens]
					g8920240	2e-89	AK000559 hypothetical protein, similar to (U06944) PRAJA1 [Mus musculus] [Homo sapiens]
					g2979531	2e-51	R33683_3 [Homo sapiens]

TABLE 6 (cont.)

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability score	Annotation
65	3	246	57	794	g5262560	3e-65	hypothetical protein [Homo sapiens]
					g10434856	4e-64	unnamed protein product [Homo sapiens]
					g930123	7e-56	zinc finger protein (583 AA) [Homo sapiens]
66	3	120	51	410	g4589566	2e-23	KIAA0961 protein [Homo sapiens]
					g456269	7e-22	zinc finger protein 30 [Mus musculus domesticus]
					g5080758	2e-20	BC331191_1 [Homo sapiens]
67	2	122	329	694	g10047297	7e-26	KIAA1611 protein [Homo sapiens]
					g8163824	2e-19	krueppel-like zinc finger protein HZF2 [Homo sapiens]
					g3329372	6e-19	DNA-binding protein [Homo sapiens]
68	3	428	132	1415	g6094684	0.0	similar to Kelch proteins; similar to BAA77027 (PID:g4650844) [Homo sapiens]
					g7242973	0.0	KIAA1309 protein [Homo sapiens]
					g7243089	0.0	KIAA1354 protein [Homo sapiens]
69	2	307	2	922	g8671168	1e-135	hypothetical protein [Homo sapiens]
					g8886025	1e-135	collapsin response mediator protein-5 [Homo sapiens]
					g8671360	1e-131	Ulip-like protein [Rattus norvegicus]
70	1	198	856	1449	g1864085	1e-103	glypican-5 [Homo sapiens]
					g3015542	1e-103	glypican-5 [Homo sapiens]
					g205800	7e-38	intestinal protein OCI-5 [Rattus norvegicus]
71	1	227	511	1191	g1155088	1e-06	zyxin [Homo sapiens]
					g1545954	1e-06	zyxin [Homo sapiens]
					g576623	2e-06	ESP-2 [Homo sapiens]
72	3	122	3	368	g7629994	4e-41	60S RIBOSOMAL PROTEIN L36 homolog [Arabidopsis thaliana]
					g3236242	5e-40	60S ribosomal protein L36 [Arabidopsis thaliana]
					g11908070	5e-40	60S ribosomal protein-like protein [Arabidopsis thaliana]

TABLE 6 (cont.)

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability score	Annotation
73	2	209	500	1126	g10435614	1e-113	unnamed protein product [Homo sapiens]
					g7243089	1e-113	KIAA1354 protein [Homo sapiens]
					g7242973	1e-107	KIAA1309 protein [Homo sapiens]
74	1	312	961	1896	g7243215	1e-157	KIAA1417 protein [Homo sapiens]
					g7263990	1e-157	dJ93K22.1 (novel protein (contains DKFZP564B116)) [Homo sapiens]
					g7302944	3e-17	CG8060 gene product [Drosophila melanogaster]
75	3	190	3	572	g10435919	6e-69	unnamed protein product [Homo sapiens]
					g3327128	3e-33	KIAA0657 protein [Homo sapiens]
					g10436504	4e-09	unnamed protein product [Homo sapiens]
76	3	295	3	887	g10436290	1e-105	unnamed protein product [Homo sapiens]
					g10436002	6e-99	unnamed protein product [Homo sapiens]
					g8489831	2e-27	ubiquitin-conjugating BIR-domain enzyme APOLLON [Homo sapiens]
77	2	288	374	1237	g3184264	5e-94	F02569_2 [Homo sapiens]
					g10435546	5e-84	unnamed protein product [Homo sapiens]
					g6653742	4e-54	7h3 protein [Homo sapiens]
78	1	294	97	978	g7670362	1e-106	unnamed protein product [Mus musculus]
					g6175860	4e-15	g1-related zinc finger protein [Mus musculus]
					g6330555	1e-13	KIAA1214 protein [Homo sapiens]
79	3	196	3	590	g3513300	3e-65	F16601_1, partial CDS [Homo sapiens]
					g3882281	3e-50	KIAA0780 protein [Homo sapiens]
					g10567164	4e-50	gene amplified in squamous cell carcinoma-1 [Homo sapiens]
80	3	745	285	2519	g2224553	0.0	KIAA0306 [Homo sapiens]
					g4210501	0.0	BC85722_1 [Homo sapiens]
					g10728201	3e-20	CG2779 gene product [Drosophila melanogaster]
81	3	256	507	1274	g6330617	1e-132	KIAA1223 protein [Homo sapiens]
					g7301689	2e-72	CG10011 gene product [Drosophila melanogaster]
					g4803678	2e-33	ankyrin (brank-2) [Homo sapiens]

TABLE 6 (cont.)

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability score	Annotation
82	1	235	841	1545	g9802433	2e-76	ACE-related carboxypeptidase ACE2 [Homo sapiens]
					g5817160	2e-76	hypothetical protein [Homo sapiens]
					g11876766	2e-76	unnamed protein product [Homo sapiens]
83	1	617	229	2079	g6665594	0.0	trp-related protein 4 truncated variant delta [Homo sapiens]
					g6665592	0.0	trp-related protein 4 truncated variant beta [Homo sapiens]
					g6665590	0.0	trp-related protein 4 [Homo sapiens]
84	3	293	735	1613	g7242977	1e-143	KIAA1311 protein [Homo sapiens]
					g912755	2e-15	B0336.3 gene product [Caenorhabditis elegans]
					g7298595	8e-12	CG10084 gene product [Drosophila melanogaster]
85	3	276	30	857	g3955100	2e-74	vacuolar adenosine triphosphatase subunit D [Mus musculus]
					g1226235	2e-74	Ac39/physophilin [Mus musculus]
					g736727	2e-74	32 kd accessory protein [Bos taurus]
86	3	355	1392	2456	g5457043	0.0	protocadherin beta 4 [Homo sapiens]
					g11142065	0.0	protocadherin beta 9 [Homo sapiens]
					g8926617	0.0	protocadherin 3H [Homo sapiens]
87	2	745	716	2950	g5457023	0.0	protocadherin alpha 9 short form protein [Homo sapiens]
					g3540157	0.0	KIAA0345-like 5 [Homo sapiens]
					g2224631	0.0	KIAA0345 [Homo sapiens]
88	2	781	50	2392	g5006248	0.0	TLR6 [Homo sapiens]
					g11596326	0.0	toll-like receptor 6 [Mus musculus]
					g5006250	0.0	TLR6 [Mus musculus]
89	2	293	1313	2191	g6164628	2e-27	SH3 and PX domain-containing protein SH3PX1 [Homo sapiens]
					g5327052	2e-27	dJ403L10.1 (SNX9 (Sorting Nexin 9)) [Homo sapiens]
					g4689258	2e-27	sorting nexin 9 [Homo sapiens]

TABLE 6 (cont.)

SEQ ID NO:	Frame	Length	Start	Stop	GI Number	Probability score	Annotation
90	1	241	214	936	g7022971	1e-62	unnamed protein product [Homo sapiens]
					g3882311	4e-15	KIAA0795 protein [Homo sapiens]
					g4539520	4e-14	dA22D12.1 (novel protein similar to Drosophila Kelch (Ring Canal protein, KEL) and a heterogenous set of other types of proteins) [Homo sapiens]

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	*PFAM hits:* Probability value= 1.0E-3 or less *Signal peptide hits:* Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

CLAIMS

What is claimed is:

1. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45,

b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45,

c) a polynucleotide sequence complementary to a),

d) a polynucleotide sequence complementary to b), and

e) an RNA equivalent of a) through d).

2. An isolated polynucleotide of claim 1, comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:1-45.

3. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 1.

4. A composition for the detection of expression of disease detection and treatment molecule polynucleotides comprising at least one of the polynucleotides of claim 1 and a detectable label.

5. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 1, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

6. A method for detecting a target polynucleotide in a sample, said target polynucleotide comprising a sequence of a polynucleotide of claim 1, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

7. A method of claim 5, wherein the probe comprises at least 30 contiguous nucleotides.

8. A method of claim 5, wherein the probe comprises at least 60 contiguous nucleotides.

9. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 1.

10. A cell transformed with a recombinant polynucleotide of claim 9.

11. A transgenic organism comprising a recombinant polynucleotide of claim 9.

12. A method for producing a disease detection and treatment molecule polypeptide, the method comprising:

a) culturing a cell under conditions suitable for expression of the disease detection and treatment molecule polypeptide, wherein said cell is transformed with a recombinant polynucleotide of claim 9, and

b) recovering the disease detection and treatment molecule polypeptide so expressed.

13. A purified disease detection and treatment molecule polypeptide (MDDT) encoded by at least one of the polynucleotides of claim 2.

14. An isolated antibody which specifically binds to a disease detection and treatment molecule polypeptide of claim 13.

15. A method of identifying a test compound which specifically binds to the disease detection and treatment molecule polypeptide of claim 13, the method comprising the steps of:

a) providing a test compound;

b) combining the disease detection and treatment molecule polypeptide with the test compound for a sufficient time and under suitable conditions for binding; and

c) detecting binding of the disease detection and treatment molecule polypeptide to the test compound, thereby identifying the test compound which specifically binds the disease detection and treatment molecule polypeptide.

16. A microarray wherein at least one element of the microarray is a polynucleotide of claim 3.

17. A method for generating a transcript image of a sample which contains polynucleotides, the method comprising the steps of:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 16 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

18. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence of claim 1, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

19. A method for assessing toxicity of a test compound, said method comprising:

a) treating a biological sample containing nucleic acids with the test compound;

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 1 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 1 or fragment thereof;

c) quantifying the amount of hybridization complex; and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

20. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, said target polynucleotide having a sequence of claim 1.

21. An array of claim 20, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

22. An array of claim 20, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide

23. An array of claim 20, which is a microarray.

24. An array of claim 20, further comprising said target polynucleotide hybridized to said first oligonucleotide or polynucleotide.

25. An array of claim 20, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

26. An array of claim 20, wherein each distinct physical location on the substrate contains multiple nucleotide molecules having the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another physical location on the substrate.

27. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) an amino acid sequence selected from the group consisting of SEQ ID NO:46-90,

b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:46-90,

c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:46-90, and

d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:46-90.

<110> INCYTE GENOMICS, INC.
PANZER, Scott R.
SPIRO, Peter A.
BANVILLE, Steven C.
SHAH, Purvi
CHALUP, Michael S.
CHANG, Simon C.
CHEN, Alice
D'SA, Steven A.
AMSHEY, Stefan
DAHL, Christopher R.
DAM, Tam C.
DANIELS, Susan E.
DUFOUR, Gerard E.
FLORES, Vincent
FONG, Willy T.
GREENAWALT, Lila B.
HILLMAN, Jennifer L.
JONES, Anissa L.
LIU, Tommy F.
ROSEBERRY, Ann M.
ROSEN, Bruce H.
RUSSO, Frank D.
STOCKDREHER, Theresa K.
DAFFO, Abel
WRIGHT, Rachel J.
YAP, Pierre E.
YU, Jimmy Y.
BRADLEY, Diana L.
BRATCHER, Shawn R.
CHEN, Wensheng
COHEN, Howard J.
HODGSON, David M.
LINCOLN, Stephen E.

<120> MOLECULES FOR DISEASE DETECTION AND TREATMENT

<130> PT-1133 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/185,213; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
60/205,324; 60/205,286
<151> 2000-02-24; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
2000-05-17; 2000-05-17

<160> 90
<170> PERL Program

<210> 1
<211> 1378
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:977683.1:2000FEB18

<220>
<221> unsure
<222> 1355
<223> a, t, c, g, or other

<400> 1
caggagatgg cggcggcggc ggctagggat cagacatggc ggcggatctg aacctggagt 60
ggatctccct gccccggtcc tggacttacg ggatcaccag gggcggccga gtcttcttca 120

tcaacgagga ggccaagagc accacctggc tgcaccccgt caccggcgag gcggtggtca 180
ccggacaccg gcggcagagc acagatttgc ctactggctg ggaagaagca tatacttttg 240
aaggtgcaag atactatata aaccataatg aaaggaaagt gacctgcaaa catccagtca 300
caggacaacc atcacaggac aattgtattt ttgtagtgaa tgaacagact gttgcaacca 360
tgacatctga agaaaagaag gaacggccaa taagtatgat aaatgaagct tctaactata 420
acgtgacttc agattatgca gtgcatccaa tgagccctgt aggcagaact tcacgagctt 480
caaaaaaagt tcataatttt ggaaagaggt caaattcaat taaaaggaat cctaatgcac 540
cggttgtcag acgaggttgg ctttataaac aggacagtac tggcatgaaa ttgtggaaga 600
aacgctggtt tgtgctttct gacctttgcc tcttttatta tagagatgag aaagaagagg 660
gtatcctggg aagcatactg ttacctagtt ttcagataag ctttgcttac cctctgaaga 720
tcacattaat cgcaaatatg cttttaaggc agcccatcca aacatgcgga cctattattt 780
ctgcactgat acaggaaagg aaatggagtt gtggatgaaa gccatgttag atgctgccct 840
agtacagaca gaacctgtga aaagagtgga caagattaca tctgaaaatg caccaactaa 900
agaaaccaat aacattccca accatagagt gctaattaaa ccagagatcc aaaacaatca 960
aaaaaacaag gaaatgagca aaattgaaga aaaaaaggca ttagaagctg aaaaatatgg 1020
atttcagaag gatggtcaag atagaccctt aacaaaaatt aatagtgtaa agctgaattc 1080
tctgccatct gaatatgaga gtgggtcagc atgccctgct cagactgtgc actacagacc 1140
aatcaacttg agcagttcag agaacaaaat agtcaatgtt agcctggcag atcttagagg 1200
tggaaatcgc cccaatacag ggcccttata cacagaggcc gatcgagtca tacagagaac 1260
aaattcaatg cagcagttgg aacagtggat taaaatccag aaggggaggg gtcatgaaga 1320
agaaaccagg ggagtaattt cttaccaaac attancaaga aatatgccaa gtcacaga 1378

<210> 2
<211> 662
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:893050.1:2000FEB18

<400> 2
gggtcttaga gtttaccttc tacttccttt agagtgtctt cgcttttctc agggcacttg 60
gaggtcctaa aactgctggt ggcacgggga gcagacctcg gctgcaaggc ccgcaagggc 120
tatgggctgc tccatacagc tgctgccagt ggccagattg aagtggtgaa gtacctgctt 180
cggatgggag cggagatcga tgaacccaat gcttttggaa acacagcttt gcacatcgcc 240
tgctacctgg gccaggatgc tgtggctatt gagctggtga atgccggagc caatgtcaac 300
cagccgaatg acaagggctt cacgccactg catgtggctg cagtctcgac caatggcgct 360
ctctgcttgg agctactggt taataatggg gctgacgtca actaccagag caaagaaggg 420
aaaagtcctc tgcacatggc tgcaatccat ggccgtttca cacgctccca gatcctcatc 480
cagaatggca gcgagattga ttgtgccgac aaatttggga acacgccact gcatgtggct 540
gctcgatatg gacacgagct gctcatcagc accctcatga ccaatggcgc agataccggc 600
cggcgtggca tccatgacat gttccccctg cacttagctg ttctctttgg attctctgac 660
tg 662

<210> 3
<211> 2764
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:980153.1:2000FEB18

<220>
<221> unsure
<222> 2663
<223> a, t, c, g, or other

<400> 3
ccccgttccc gattcatgta gtagcggctg tattgcagcc gcctgccgaa ctgacccggg 60
tctggggact ggcccctctg gcgccgttcg gtttctctta ttgccttcac tgaggatgag 120
tccctttgtg gctctatgtg gaccctgcgg aatccaccgg cgcagtttca tctagcgact 180
ggtcaccctt gcaattatgg atatttaaaa gggtcagaca gtgtggaggg ggagttcccc 240
tcctcactcc cccttggtgc ttgactccag gaataattta taaactgtgg aatttttta 300
aatgaagaac ttgtatttga tatgaacttt atagagctat ttataatttt tttgatttaa 360
gtgccaaaaa aattgtataa agatatatag ttttatacta ttgtcaggag gatttaaatt 420

atcctaaaaa ggtaatttat tctctgtaac ttcctcaata gcacctttgt gtcctggctt 480
tttcattttt taaaattagt tttcacgatt ctgaagtaag tggtataaaa acagttagga 540
tgagttcacc catgcctgac tgcacatcaa agtgtcgatc cctgaagcat gctttggaag 600
tcccttctgt ggtaacaaag gggagcgaaa acccgattaa ggcccttctc tccacgtcat 660
tgttacaaag ctgccactat caaggatgtt tttggcagga atgccctcca cccctgtttc 720
ctcctcgtgg agaagaaagg agtgttagat tggcttattc agaaaggagt ggatctgttg 780
gtgaaagaca aagagtctgg atggacagca ttgcacagaa gcatttttta tggacatatt 840
gattgtgttt ggtctctatt gaagcatggt gttagtctgt atattcaaga taaagaaggc 900
ttgtcagctt tggatcttgt aatgaaggat agaccaactc atgtagtatt caagaatact 960
gatcctacag atgtttatac ttggggcgat aatacaaatt ttaccctggg tcatggaagc 1020
cagaatagca aacatcatcc agagttggtg gatctgttct ccaggagtgg gatttatatc 1080
aagcaggtgg tgctttgtaa atttcactcc gtgtttctgt ctcagaaagg gcaggtttat 1140
acctgtggtc atggtcctgg agggcgatta ggacatggag atgaacagac atgcttggtc 1200
cctcggcttg tggaaggact gaatggtcat aattgttccc aagtggcagc tgctaaggat 1260
catactgttg tattaactga agatggatgt gtttatacat ttggtctaaa cattttcat 1320
caattaggaa ttattccacc gccttccagt tgtaatgtac ccagacagat acaggcaaaa 1380
tatctgaaag gaaggacaat cattggcgtt gcagcaggca ggtttcatac agtcctatgg 1440
actagagaag ctgtttacac tatgggacta aatggtggac aactgggttg tttgctagat 1500
cccaatggag aaaagtgtgt aactgctcct cgtcaggtct ctgcccttca ccataaagac 1560
attgctctgt ctttggttgc tgcaagtgat ggagctacag tctgtgttac cacaagggga 1620
gatatttact tacttgcaga ctatcagtgc aagaagatgg cttctaaaca gttgaacttg 1680
aaaaaagttc ttgtgtctgg gggtcatatg gaatacaagg ttgatcctga acatttgaaa 1740
gaaaatgggg gtcaaaaaat ttgcattctt gcaatggatg gagctggaag ggtgttttgc 1800
tggagatcag tcaacagttc tctgaagcag tgtcgatggg cctatccacg tcaggtcttc 1860
atttctgata ttgctttaaa tagaaatgaa attctatttg ttacgcaaga tggagaagga 1920
tttagaggga gatggtttga agagaaaaga aagagttctg aaaagaaaga gattttatca 1980
aaccttcaca attcctcatc agatgtgtct tatgtctctg atataaatag tgtgtatgaa 2040
agaattcgac ttgagaaact tacctttgca catagagctg ttagtgtcag cacagatcca 2100
agtggatgca actttgcaat cctgcagtca gatcctaaaa caagccttta tgaaattcca 2160
gctgtgtcct catcatcctt ttttgaagag tttggcaaac tgttgaggga agcagatgaa 2220
atggacagca ttcatgatgt gacatttcaa gttggcaata gactcttccc tgcacataaa 2280
tatattttgg cagtgcattc tgatttttt cagaaattgt ttctttcaga tggtaatact 2340
tcagaattta cagatattta ccagaaagat gaagattctg cagggtgcca tctctttgtg 2400
gtagagaagg ttcatcctga catgtttgaa tacctttac aatttatata cacagatact 2460
tgtgactttt taactcatgg cttcaaacca agaatacact taaacaaaaa cccagaagaa 2520
tatcagggaa ctctgaattc tcatttgaat aaagtgaatt tccatgaaga tgataaccag 2580
aagtctgcat ttgaagttta caaaagtaat caagctcaaa cagttagtga gaggcagaag 2640
agcaaaccta aatcttgtaa aanaggaaaa aatattaggg aagatgatcc tgtaagaatg 2700
ttgcaaactg ttgcaaagaa attcgacttc agtaatttga gtagtaggtt agatggagtc 2760
agat 2764

<210> 4
<211> 388
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:350398.1:2000FEB18

<220>
<221> unsure
<222> 125
<223> a, t, c, g, or other

<400> 4
cccttctacg tccgctgcat caagcccaat gaggacaagg tagctgggaa gctggatgag 60
aaccactgtc gccaccaggt cgcatacctg gggctgctgg agaatgtgag ggtccgcagg 120
gctgnttcgc ttcccgccag ccctactctc gattcctgct caggtactgg cacctgacac 180
ccatcactcc atgggccata gtccctgtgt ggagtccaag gggtaggagc agagggtccc 240
caaacagcac gtcgcaaaca tcgatacaag caggaaccag cacgctgctg gcctcaagac 300
accaaaatat ctgggaagac atgtgtgtga gcacatgcat gtggggacat acaggtggga 360
acatgggtat gagggctgtg tgaggaca 388

<210> 5
<211> 2364
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:475551.1:2000FEB18

<220>
<221> unsure
<222> 424-425
<223> a, t, c, g, or other

<400> 5
gtctgcaggg ccagagcggg gcagacatgg acaagcgggt gaagaagctt cccctcatgg 60
ctctgtccac cacgatggct gagagcttca aggagctgga ccctgattcc agcatgggga 120
aggccttgga gatgagctgt gccatccaga atcagctggc ccgcatcctg gccgagtttg 180
agatgaccct ggagagggac gtcctgcagc cactcagcag gctgagtgag gaggagctgc 240
cagccatcct caaacacaag aaaagcctcc agaagctcgt gtccgactgg aacacactca 300
agaacaggct cagtcaggca accaagaatt caggcagcag tcaaggccta ggaggcagcc 360
cgggtagtca cagccatacg accatggcca acaaggtgga gacgctgttc tactgcagca 420
ggtnntcacc caggaaagtg gagcaatgca gggacgagta cttggctgac ctgtaccact 480
ttgttaccaa ggaggactcc tatgccaact acttcattcg tctcctggag attcaggccg 540
attaccatcg caggtcactg agctcgctgg acacagccct ggctgagctg agggagaacc 600
acggccaagc agaccactcc ccttcgatga cagccaccca cttccccagg gtgtatgggg 660
tgtcgctggc aacccacctg caagagctgg gccgggagat tgccctgccc atcgaggcct 720
gcgtcatgat gctgctttct gagggcatga aggaagaggg tctcttccgt ctggctgctg 780
gggcctcggt gctgaagcgt ctcaagcaga caatggcctc ggacccccac agcctggagg 840
agttctgctc cgacccgcac gctgtggcag gtgccctcaa gtcctatctg cgggagctgc 900
cagagcctct gatgaccttc gacctctatg atgactggat gagggcagcc agcctgaagg 960
agccaggggc ccggctgcag gccctccaag aggtgtgcag ccgcctaccc cccgagaacc 1020
tcagcaacct caggtacctg atgaagttcc tggcacggct ggccgaggag caggaggtga 1080
acaagatgac acccagcaac atcgccatag tcctgggacc caacttgctg tggccacctg 1140
agaaagaagg ggaccaggcc cagctggatg cagcctccgt gtcttccatc caggtggtgg 1200
gcgtcgtcga ggcgctgatc cagagcgcag acaccctctt ccctggagac atcaacttca 1260
acgtgtcagg cctcttctca gctgttaccc tccaggacac agtcagtgac aggctggcct 1320
ctgaggaact tccgtccact gccgtgccca ccccagccac caccccggct ccggctccgg 1380
ctccagctcc agctccggcc ccagccttgg cttcagcagc taccaaggaa aggacagagt 1440
ctgaggtgcc tcccagacca gcctccccca aggtcaccag gagtcccccg gagacagctg 1500
ccccagtgga ggacatggct cggaggacca agcgcccggc gccagccccgg cccaccatgc 1560
cgcccccccа ggtctccggc tcccgctcct cccctccagc cccgcccttg ccccctggct 1620
ctggcagccc tgggaccccc caagccctgc cccgacgtct ggttggcagc agcctccgag 1680
cccccacagt gccacccccg ttacccccca caccccctca gcctgccccgg cgccaaagcc 1740
ggcgttcacc agcctccccc agcccggcct ccccaggtcc agcctccccc agcccagtct 1800
ctttgagtaa ccctgcacag gtggacctgg gggctgccac agcagaggga ggagcccctg 1860
aggctatcag tggggtcccc actcccccag ctatcccccc tcagccccgc cccaggagcc 1920
ttgcctcaga gaccaactga gtggctggtt tctccctaag cagccctcag caccccctcc 1980
ctccccacct ggccctccca ggacagctct cgcccccccac aaaggggcat gggcctccag 2040
cctttgccca caagtgcctc agtgcccact gggtcggccc ccatggccag gagggctcag 2100
gacaatcctc tatttcctga ccttttcctc gtccaccctg ggcttgggga ccccccccacc 2160
ggactctcca ctctccggca ggtcctaggg gagccaccgg aaggaaggag aggtttgcct 2220
gctcctacgg gactgattct tctcttgccg acatgttttt tgtaaggctg gtaaataaat 2280
tattttggac aaaactggag cagctgccca aatgatagtt ttatttctg tccttgaaat 2340
aaagaagcca attttataaa gggg 2364

<210> 6
<211> 1801
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:481407.2:2000FEB18

<400> 6
cgccttcggg gcccggatct caaacagtcg ggaagaagca ccgtggctgc tattatctgc 60
tctccgcgcc tgacccctcc caggactcgt gatgccaagg ccgctgcgag cggctacgaa 120
gagtcggggt tgagccccag ctgagccgag ggctcgcact cttctggtct cccaggccca 180
acccacctga agaaatgagt ggtggattgg ctccaagtaa gagcacagtg tatgtatcca 240

acttgccttt ttccctgaca aacaatgact tgtaccggat attttccaag tatggcaaag 300
ttgtaaaggt taccatcatg aaagataaag ataccaggaa gagtaaaggg gttgcattta 360
ttttattttt ggataaagac tctgcacaaa actgtaccag ggcaataaac aacaaacagt 420
tatttggtag agtgataaaa gcaagcattg ctattgacaa tggaagagca gctgagttca 480
tccgaaggcg aaactacttt gataaatcta agtgttatga atgtggggaa agtggacact 540
taagttatgc ctgtccgaaa aatatgctcg gagaacgtga gcctccaaag aagaaagaaa 600
aaaaagaaaa aaaagaaagc tcctgaacca gaagaagaaa ttgaggaagt agaagaaagt 660
gaagatgaag gggaggatcc tgctcttgac agcctcagtc aggccatagc attccagcaa 720
gccaaaattg aagaagaaca aaaaaaatgg aaacccagtt caggagtccc ctcaacatca 780
gatgattcaa gacgcccaag gataaagaaa agcacatatt tcagtgatga ggaagaactt 840
agtgattaaa atcttgcccc agcacagtaa taaaaatcaa gatttgttag taacaatctt 900
gaagagctaa ttttaataaa aataagaaaa attaatacta tcatgttaat actattattg 960
tcatcccaag aaaaaagata ttttaaaaat ttatttgaaa agttcattat aagggcttta 1020
ttcatgcctg atttgtttac atgaggactt ctgaaattaa tccttaaaac aaacttcctg 1080
aagaccgaaa agttgaatga tttattgtta cttatattaa taaacttttc aagagaattt 1140
tgtctttaaa tatgggtgtt ttgtcatcat atttcttgta gctttatccc aatctggata 1200
aattgtaaat acctataaaa taaattataa atacctataa aatataaagt aacatagctc 1260
taaaaggctt aaaatcaaac acaggtgtta tttgtctgcc ctacccatag caccaaattc 1320
cattccctag aagaaactac ttacatgtgc acacatgtag ctaaataaag tacatatatt 1380
tcagtcactt aataaaagaa taaaggagaa ttattcaatc tcttatactt ctcctacctt 1440
cctcaatatt cccaatgtgg ctgtattaaa aatttggggg aatccatata tatcttttt 1500
aataaccaag taaatacttt tcatttctga gccaaggaat atgctatgat tacgtttttt 1560
cctagagtta ataattgtct attttttttc catgtattgt ctttgtattt atgactaaat 1620
cttcccattc tgtctgcagg tgggtatatg gtaatgggat tagagagcct ttaattttct 1680
gctttgtata tttctatatt gtttaacttt gtaagaatgc ccattacttt tttaactagt 1740
aaaagcaata gaaataagtt aatactatca tagtaatatt attattgtca tcccgaggcg 1800
g 1801

<210> 7
<211> 730
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:443580.1:2000FEB01

<220>
<221> unsure
<222> 44
<223> a, t, c, g, or other

<400> 7
ggaggtgaga tattttggtc cccaggagaa ctggctcagg tctncaagtt cccatccggg 60
atgactggaa agggttagga aacctctctg aggtctggtc agattccaac cctggacagc 120
agtgaacaca acctttcccc tgagccactg gaattggaca gaatgcccca ttctcctctg 180
atctccattc ctcatgtgtg gtgtcaccca gaagaggagg aaagaatgca tgatgaactt 240
ctacaagcag tatccaaggg gccggtgatg ttcagggatg tttccataga cttctctcaa 300
gaggaatggg aatgcctgga cgctgatcag atgaatttat acaaagaagt gatgttggag 360
aatttcagca acctggtttc agtgggactt tccaattcta agccagctgt gatctcctta 420
ttggaacaag gaaaagagcc ctggatggtt gatagagagc tgactagagg cctgtgttca 480
gatctggaat caatgtgtga gaccaaaata ttatctctaa agaagagaca tttcagtcaa 540
gtaataatta cccgtgaaga catgtctact tttattcagc ccacatttct tattccacct 600
caaaaaacta tgagtgaaga gaaaccatgg gaatgtaaga tatgtggaaa gacctttaat 660
caaaactcac aatttatcca acatcagaga attcattttg gtgaaaaaca ctatgaatct 720
aaggaaaaaa 730

<210> 8
<211> 457
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:803015.1:2000FEB01

<400> 8

gcgcgggctg cggctgggat ccggtctttc cagccccgag agggacctgg ttcctctgcc 60
caggcttctg tcactctgtc acctacgcta tgccctgctg tagtcacagg aggtgtagag 120
aggaccccgg gacatctgaa agccaggaaa tggacccagt ggcctttgat gatgttgctg 180
tgaacttcac ccaggaggag tgggctttgc tggatatttc ccagaggaaa ctctacaagg 240
aagtgatgct ggaaactttc aggaacctga cctctgtagg aaaaagttgg aaagaccaga 300
acattgaata tgagtaccaa aaccccagga gaaacttcag gagtctcata gaaaagaaag 360
tcaatgaaat taaagatgac agtcattgtg gagaaacttt tacccaggtt ccagatgaca 420
ggctgaactt ccaggagaag aaagcttctc ctgaaat 457

<210> 9
<211> 582
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:027410.3:2000MAY19

<400> 9
ggcacaccga ggctcggccg ccccgccgcg agtcctggat cagtgacatt cgagcaggaa 60
ccgccccttc atgcaggaac cacatcaaat caagctgcag cttgatcgcc ttcaactccg 120
accgtcctgg tgtactgggc attgtgcctc tgcaaggcca aggagaggac aagcgacgcg 180
tggcccacct gggctgccat tcagacctag tcaccgactt ggacttctcg ccctttgatg 240
acttcctcct ggccacaggc tcggctgaca ggacggtaaa actctggcga ctgccagggc 300
ctggccaggc cctgccctca gcacccgggg tggtgctggg ccccgaggac ctcccagtgg 360
aggtactgca gttccacccc acctctgacg gcattctgag ctggcagccc atggggacct 420
ggtgcagagc gccgtctgga gccgagatgg agccctggtg ggcacggcgt gcaaggacaa 480
gcagctgcgg atctttgacc ccagaacaaa gccgcgggcc tctcagagca cgcaggccca 540
tgagaacagc agggatagcc ggctggcatg gatgggcacc tg 582

<210> 10
<211> 848
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:171377.1:2000MAY19

<400> 10
agcggccgca gcctctgaga gcacgaacag cagcgccccc gcgtcccagc cagccagcca 60
gccagactgg actccggccc accgacggcc gctcgcgctc cggccccgct cgcctgctct 120
gccccggacc tgcagctccc cgctcccccg ccgtgtccgc cgcctcccgg ccagagagcc 180
aagccccccac gccgcgccca gccgtcgccg cgccgagcat gtcctcgacc gagaggcgcc 240
cggcgggacg gcgggacagg tcgccgcgcc agcaggtgga ccgcctactc gtggggctgc 300
gctggcggcg gctggaggag ccgctgggct tcatcaaagt tctccagtgg ctctttgcta 360
ttttcgcctt cgggtcctgt ggctcctaca gcggggagac aggagcaatg gttcgctgca 420
acaacgaagc caaggacgtg agctccatca tcgttgcatt tggctatccc tgcaggttgc 480
accggatcca atatgagatg cccctctgcg atgaagagtc cagctccaag accatgcacc 540
tcatggggga cttctctgca cccgccgagt tcttcgtgac ccttggcatc ttttccttct 600
tctataccat ggctgcccta gttatctacc tgcgcttcca caacctctac acagagaaca 660
aacgcttccc gctggtggac ttctgtgtga ctgtctcctt caccttcttc tggctggtag 720
ctgcagctgc ctggggcaag ggcctgaccg atgtcaaggg ggccacacga ccatccagct 780
tgacagcagc catgtcagtg tgccatggag aggaagcagt gtgcagtgcc ggggccacgc 840
cctctatg 848

<210> 11
<211> 636
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:352559.1:2000MAY19

<400> 11
tgtagtttcc tcaactactg cctcagctct acaatcccag agtaaagctc ttctccaaat 60

gaagagccag gaagaggtag aggtggcagg aattaaactt tgtaaagcca tgtccctggg 120
ttcactgact ttcacagatg tggccataga cttttcccaa gatgaatggg agtggctgaa 180
tcttgctcag agaagtttgt acaagaaggt gatgttagaa aactacagga acctagtttc 240
agtgggtctt tgcatttcta aaccagatgt gatctcctta ctggagcaag agaaagaccc 300
ttgggtgata aaaggaggga tgaacagagg cctgtgccca gacttggagt gtgtgtgggt 360
gaccaaatca ttatctttaa accaggatat ttatgaagaa aaattacccc cggcaatcat 420
aatggaaaga cttaaaagct atgaccttga atgttcaaca ttagggaaaa actggaaatg 480
tgaagacttg tttgagaggg agcttgtaaa ccagaagaca cattttaggc aagagaccat 540
cactcatata gatactctta ttgaaaaaag agatcactct aacaaatctg ggacagtttt 600
tcatctgaat acattatctt atataaaaca gatttt 636

<210> 12
<211> 2110
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:247384.1:2000MAY19

<400> 12
ccaggagaag gaagccaaca ggatccgacc cggtgttttg tgacaaaggc aagaccccca 60
ggtctactta gagcaaagtt agtagaggag gcagctaggc gtggctctca ttccttccca 120
cagaatggat tataagtcga gcctgatcca ggatgggaat cccatggaga acttggagaa 180
gcagctgatc tgccctatct gcctggagat gtttaccaag ccagtggtca tcttgccgtg 240
ccagcacaac ctgtgccgga agtgtgccaa tgatattttc caggcctcta acccgtattt 300
gcccacaaga ggaggtacca ccatggcatc agggggccga ttccgctgcc catcctgtag 360
acatgaagtg gttttggata gacatggggt atatggactt cagaggaacc tgctggtgga 420
gaacatcatc gacatctaca aacaggagtg ctccagtcgg ccgctgcaga agggcagtca 480
ccccatgtgc aaggagcacg aagatgagaa aatcaacatc tactgtctca cgtgtgaggt 540
gcccacctgc tccatgtgca aggtgtttgg gatccacaag gcctgcgagg tggccccatt 600
gcagagtgtc ttccagggac aaaagactga actgaataac tgtatctcca tgctggtggc 660
ggggaatgac cgtgtgcaga ccatcatcac tcagctggag gattcccgtc gagtgaccaa 720
ggagaacagt caccaggtaa aggaagagct gagccagaag tttgacacgt tgtatgccat 780
cctggatgag aagaaaagtg agttgctgca gcggatcacg caggagcagg agaaaaagct 840
tagcttcatc gaggccctca tccagcagta ccaggagcag ctggacaagt ccacaaagct 900
ggtggaaact gccatccagt ccctggacga gcctgggggа gccaccttcc tcttgactgc 960
caagcaactc atcaaaagca ttgtggaagc ttccaagggc tgccagctgg ggaagacaga 1020
gcagggcttt gagaacatgg acttctttac tttggattta gagcacatag cagacgccct 1080
gagagccatt gactttggga cagatgagga agaggaagaa ttcattgaag aagaagatca 1140
ggaagaggaa gagtccacag aagggaagga agaaggacac cagtaaggag ctggatgaat 1200
gagaggcccc cagatgcaga gagactggag agggtgggga ggggcccagc ggccttggtg 1260
acaggcccag ggtgggaggg gtcggggccc ctggaggggc aatggggagg tgatgtcttc 1320
tctctgctca gagagcaggg actagggtag gaccctcacc gctgcgtcca gcagacactg 1380
aaccagaatt ggaaacgtgc ttgaaacaat cacacaggac acttttctac attggtgcaa 1440
aatggaatat tttgtacatt tttaaaatgt gatttttgta tatacttgta tatgtatgcc 1500
aatttggtgc tttttgtaaa ggaacttttg tataataatg cctggtcatt gggtgacctg 1560
cgattgtcag aaagagggga aggaagccag gttgatacag ctgcccactt cctttcctga 1620
gcaggaggat ggggtagcac tcacagggac gatgtgctgt atttcagtgt ctatcccaga 1680
catacggggt ggtaactgag tttgtgttat atgttgtttt aataaatgca caatgctctc 1740
ttcctgttct tcaaaggagc cggggtttca ttcagccttt ttttcctgga gatgagggtt 1800
gagtgtgaat gaacaggacc cctggtagga ggcaatggca gggctaggct taggtcccag 1860
taaaggagtt ctcgacacca ccatttccca atgtggactc catggaaagc cagccctgag 1920
ctggtccttc aagaacaggt tcaatgtgtt gttgctctgg ttctccagaa aacagagcct 1980
gaggcaaaat ttaaatgctt tagttgcagg ttatagggac ttccccgtgc tcactgaagg 2040
ctcactgaag gctcactgaa aatcatcaag aagaggcaga ttaaggctgg gcaggtgcag 2100
tggttcatgc 2110

<210> 13
<211> 2375
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:403872.1:2000MAY19

<220>
<221> unsure
<222> 1233
<223> a, t, c, g, or other

<400> 13
gcagcgccag gaggaggcag cggaggaagc agagcgcggg atgggcgccc agcggcatct 60
gtgatcccgc gcacctccgc cccacgggcg cgcgcacaaa cacggacaca cacatacaca 120
cactcgcgca cacactcgca caaacacaca ctcgtacacg cccgcgccgc tcgctcgccg 180
gcttgctctc ccacgcaagc ggaatgcagc agcgcctgga gagcgtgtct cggaccgccg 240
cctgaatgta cctcgctccc gggagccgga cggcccagta gggcgcactg gaggacgctc 300
cgctgcggga gcctggacag tttttgacgg tgcagtcttg ctatatggtg tgagaaatgg 360
ctgtaggaaa caacactcaa cgaagttatt ccatcatccc gtgttttata tttgttgagc 420
ttgtcatcat ggctgggaca gtgctgcttg cctactactt cgaatgcact gacacttttc 480
aggtgcatat ccaaggattc ttctgtcagg acggagactt aatgaagcct tacccaggga 540
cagaggaaga aagcttcatc acccctctgg tgctctattg tgtgctggct gccaccccaa 600
ctgctattat ttttattggt gagatatcca tgtatttcat aaaatcaaca agagaatccc 660
tgattgctca ggagaaaaca attctgaccg gagaatgctg ttacctgaac cccttacttc 720
gaaggatcat aagattcaca ggggtgtttg catttggact ttttgctact gacatttttg 780
taaacgccgg acaagtggtc actgggcact taacgccata cttcctgact gtgtgcaagc 840
caaactacac cagtgcagac tgccaagcgc accaccagtt tataaacaat gggaacattt 900
gtactgggga cctggaagtg atagaaaagg ctcggagatc ctttccctcc aaacacgctg 960
ctctgagcat ttactccgcc ttatatgcca cgatgtatat tacaagcaca atcaagacga 1020
agagcagtcg actggccaag ccggtgctgt gcctcggaac tctctgcaca gccttcctga 1080
caggcctcaa ccgggtctct gagtatcgga accactgctc ggacgtgatt gctggtttca 1140
tcctgggcac tgcagtggcc ctgtttctgg gaatgtgtgt ggttcataac tttaaaggaa 1200
cgcaaggatc tccttccaaa cccaagcctg agnatccccg tggagtaccc ctaatggctt 1260
tcccaaggat agaaagccct ctggaaacct taagtgcaca gaatcactct gcgtccatga 1320
ccgaagttac ctgagacgac tgatgtgtca caagctgttt tttaaaatca tcttccaatt 1380
ctatacttca aaacacacag ttgctcaatg tcaaactgtg atgacaaata ttacgtttat 1440
ctagttagaa gctaatgttt tgtacatttt ttgtatgagg aagtgatgta gcttgccctg 1500
attttttttt ttttttttg gtcagcttta atatatttat gccagaattt taaaaccaac 1560
aaaattttct tgttcaagcg tgcattgaag aaccacattt attcaatggt tgacgttgtt 1620
ttgtgatatt tgtacacaaa ttttcttttc tcagttttat aaacacagaa gtaaatataa 1680
caattcactt taaacttta ttaccacagt tgctgcctcc tccagaattt ttgaatttta 1740
ataaaaggca aacttttgag ctgcaggaag gacaatgttg gttaataata aatctcaaag 1800
tcaattgtag aaaaaaaatt gtcttcaaaa agaatgttgc actctgatct cttaacaaat 1860
tgttacgttc aaagtttaaa gtgatatatt aacaaagtca cctagttata caaacaattg 1920
tcagagaatt ctggatttgg agggtattgg ggttatatga ttctttctta gataatggcc 1980
tctactaaat aactcaagat ctttctggaa tgtcttctgg caggcaggtg ccactgtcag 2040
cttttctcca aaaagcagcc aacatcagcc tcccctgtca actcaacagt tttgtatctc 2100
atattatatg gactttatat gaaaatgaat attttacagt ttgcacagta ttattttaca 2160
gaaaaggaat cagagaatct acaacatagg gccccagaac aacagtttca ctttgtggct 2220
tttaattatt ctagaatttt aactgcatct cattttttcta gcatggtgag aactaatatg 2280
taactccttt gattgaagga gctcttttgt ccgtacctat cagaatgttt tcttgacact 2340
tccatgttgg ctcttctcag ctttttttgt acata 2375

<210> 14
<211> 537
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1135213.1:2000MAY19

<400> 14
ggacccgcga gcggagcggc gcgtgggtcg gttgcggtcg gccccggcag gatgggaagg 60
ccattgtgac tatgtggtga ttacagttgt cttactactg agtttcctac tgaaatcatg 120
gaggagaaac agcagattat attggctaat caagatggtg gaacagtggc aggagcagca 180
cctaccttct ttgtcatctt aaagcagcca ggaaatggca aaactgatca aggaattttg 240
gttactaatc aggatgcctg tgctttggct agtagtgtgt catcaccagt aaaatctaaa 300
gggaagattt gccttccagc tgattgtact gtgggtggaa tcactgttac cctcgataac 360
aatagtatgt ggaatgagtt ctatcatcga agcacagaga tgattctgac caagcaagga 420
agacgcatgt ttccttactg tcgttattgg ataacaggtt tagattcaaa tttgaagtat 480
attcttgtca tggatatatc tcctgtggat aaccatcgtt ataagtggaa tggtcgt 537

<210> 15
<211> 1433
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:474284.2:2000MAY19

<400> 15
ggcctgcccc ggccccctgc ccgcggcgcc atggcggaga attggaagaa ctgcttcgag 60
gaggagctca tctgccctat ctgcctgcac gttttcgtgg agccagtgca gctgccgtgc 120
aaacacaact tctgccgggg ctgcatcggc gaggcgtggg ccaaggacag cggcctcgta 180
cgctgcccag agtgcaacca ggcctacaac cagaagccgg gcctggagaa gaacctgaag 240
ctcaccaaca tcgtggagaa gttcaatgcc ctgcacgtgg agaagccgcc ggcggcgctg 300
cactgcgtgt tctgccgccg cggccccccc gctgcccgcg cagaaggtct gcctgcgctg 360
cgaggcgccc tgctgccagt cccacgtgca gacgcacctg cagcagccct ccaccgcccg 420
cgggcacctc ctggtggagg cggacgacgt gcgggcctgg agctgcccgc agcacaacgc 480
ctaccgcctc taccactgcg aggccgagca ggtggccgtg tgccagtact gctgctacta 540
cagcggcgcg catcagggac actcggtgtg cgacgtggag atccgaagga atgaaatccg 600
gaagatgctc atgaagcagc aggaccggct ggaggagcga gagcaggaca ttgaggacca 660
gctgtacaaa ctcgagtcag acaagcgcct ggtggaggag aaagtgaacc aactgaagga 720
ggaagttcgg ctgcagtacg agaagctgca ccagctgctg gacgaggacc tgcggcagac 780
agtggaggtc ctagacaagg cccaggccaa gttctgcagc gagaacgcag cgcaggcgct 840
gcacctcggg gagcgcatgc aggaggccaa gaagctgctg ggctccctgc agctgctctt 900
tgataagacg gaggatgtca gcttcatgaa gaacaccaag tctgtgaaaa tcctgatgga 960
cagcagatgc cccgtccact ggccccagga cccagacctg cacgagcagc agcctttccc 1020
ccactaagat cggccacctg aactccaagc tcttcctgaa cgaagtggcc aagaaggaga 1080
agcagctgcg gaaaatgcta gaaggccccc tcagcacgcc ggtgcccttc ctgcagagtg 1140
tcccccctgta cccttgcggc gtgagcagct ctggggcgga aaagcgcaag cactcaacgg 1200
ccttcccaga ggccagtttc ctagagacgt cgtcgggccc tgtgggcggc cagtacgggg 1260
cggcgggcac agccagcggt gagggccagt ctgggcagcc cctggggccc tgcagctcca 1320
cgcagcactt ggtggccctg ccgggcggcg cccaaccagt gcactcaagc cccgtgttcc 1380
ccccatcgca gtatcccaat ggctccgcgc ccagcagccc atgctccccc agt 1433

<210> 16
<211> 654
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:342147.1:2000MAY19

<400> 16
cgaattgggc ccctagatgt ttgctcgagc ggcggccgca gtgtgctgga aagggacaaa 60
gacttgtaac tggagaaata gtttgtaagg gagatttttc ttcctctacc cacacctttc 120
aaggcaggga gcaatgaaag acaaacctgt actgttcacc atatttcatt gattgcaata 180
ggagtattga ggtcactttt atattgtcct ggatagtatg tagttacgcg gtttgtaaag 240
agaggaatgg gatgggggc tgtgagaagg aagaattagt ggtcgatttc ggaggagcag 300
gatggagatc cctgtgcctg tgcagccgtc ttggctgcgc cgcgcctcgg ccccgttgcc 360
cggactttcg gcgcccggac gcctctttga ccagcgcttc ggcgaggggc tgctggaggc 420
cgagctggct gcgctctgcc ccaccacgct cgccccctac tacctgcgcg cacccagcgt 480
ggcgctgccc gtcgcccagg tgccgacgga ccccggcccc ttttcggtgc tgctagacgt 540
gaagcacttc tcgccggagg acattgctgt caaggtggtg ggcgaacacg tggaggtgca 600
cgcgcgccac gaggagcgcc cggatgagca cggattcgtc gcgcgcgagt tcca 654

<210> 17
<211> 1651
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1097300.1:2000MAY19

<400> 17

gccgccgagg aggaggccct gctggtttct gtgcgggctc ttgtccagga tggtgaagct 60
gttcatcgga aacctgcccc gggaggctac agagcaggag attcgctcac tcttcgagca 120
gtatgggaag gtgctggaat gtgacatcat taagaattac ggctttgtgc acatagaaga 180
caagacggca gctgaggatg ccatacgcaa cctgcaccac cacaagcctc atggggtgaa 240
catcaacgcg gaagccagca agaacaagag caaagcccca accaagttac acgtgggcaa 300
catcagcccc acctgcacca accaagagct tcgagccaag tttgaggagc acggtccggc 360
catcgaatgt gacatcgcga aagactatgc cttcgcacac atggagcggg cagaggacgc 420
agcggaggcc atcaggggcc tcgacaacac agagtttcaa ggtgaactgc tctgggcctg 480
ggtagtagcg ccgagtgggg tctagctcaa aacaggcaag aacacaagac tatagaactt 540
gctgggtggt ctcttccatt ctgttttagc tggaaataat agattatgtt taccgctctt 600
aagcataatt taccccctggg gaagcaaaca cttcctcttt tcaggtttgc taagatgttg 660
ctcaccgact gcatagaatc acaaactgtg ggttacttta ccctgcggga ttcttgcatt 720
gattcgagtg ctgttggaag tgtaatctgc ttggggaaac gagtacctca tgagagaagg 780
gaggataaag gtccgtggct tacctgcttc tttggtgatg atcaggaagc cttatatttg 840
agggtttaag tgcttaagat ttatattctt tactgctttg ggtggatact ggtgggaaag 900
aagaaaaaag acatctagag gaagccctat attataaatc tgggtggcaa gtctggatct 960
gcgggagtat ctttttgttg atcaaagttg tgcagtctct tcaagcagag tcaaaaaaac 1020
atgccatgga gtgttctgct ccacctgttc atttcaccct cagaaaagga aatttctaaa 1080
tatatcagac tcaatgggaa tgatggtccc gcttctgaag aaatttcagt acaagcatcg 1140
tagagcatat catactattt ataccgataa taaaggtaca tatgttgtca ttaataccac 1200
aagaggttgt cagaagactc tagaactgtg ctaatatggt aaccacatgc ggcttagtaa 1260
attgaaatta acagattaga taaaatttaa aattcagttt ttcaagtgta taccagacac 1320
gtttcaagca ctcagtagtc atgaggcctg tggctaccgt attaatagag acacagaaca 1380
tttccatcat catagaacat tcttttggat agcactgttc tacaagtgtt ttgttaacag 1440
tatcgtcttg gacctcatgt tcatagccac ttttgtggtt cctaagtcaa caccttttt 1500
gccctgagtg tcattaaagg ggttgttaag aagtactttt gggtcttcta ttaaaactaa 1560
aaaacaaaat gagaaaaata atgggagaag aggaaaagtt gaccagagaa gggtaagaaa 1620
gtttgcatag tggagatggg tagaggagca c 1651

<210> 18
<211> 1870
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:444850.9:2000MAY19

<220>
<221> unsure
<222> 1865, 1867
<223> a, t, c, g, or other

<400> 18
ggctctgaag ccattacaaa ggttgcttaa cttctaatta tttgatcact gaggaaaatc 60
cagaaagcta cacaacactg aaggggtgaa ataaaagtcc agcgatccag cgaaagaaaa 120
gagaagtgac agaaacaact ttacctggac tgaagataaa agcacagaca agagaacaat 180
gccctggaca tggctccaga gatccacatg acaggcccaa tgtgcctcat tgagaacact 240
aatggggaac tggtggcgaa tccagaagct ctgaaaatcc tgtctgccat tacacagcct 300
gtggtggtgg tggcaattgt gggcctctac cgcacaggaa aatcctacct gatgaacaag 360
ctagctggga agaataaggg cttctctctg ggctccacag tgaaatctca caccaaagga 420
atctggatgt ggtgtgtgcc tcaccccaaa aagccagaac acaccttagt cctgcttgac 480
actgagggcc tgggagatgt aaagaagggt gacaaccaga atgactcctg gatcttcacc 540
ctggccgtcc tcctgagcag cactctcgtg tacaatagca tgggaaccat caaccagcag 600
gctatggacc aactgtacta tgtgacagag ctgacacatc gaatccgatc aaaatcctca 660
cctgatgaga atgagaatga ggattcagct gactttgtga gcttcttccc agattttgtg 720
tggacactga gagatttctc cctggacttg gaagcagatg gacaacccct cacaccagat 780
gagtacctgg agtattccct gaagctaacg caaggtacca gtcaaaaaga taaaaatttt 840
aatctgcccc aactctgtat ctggaagttc ttcccaaaga aaaaatgttt tgtcttcgat 900
ctgcccattc accgcaggaa gcttgcccag cttgagaaac tacaagatga agagctggac 960
cctgaatttg tgcaacaagt agcagacttc tgttcctaca tctttagcaa ttccaaaact 1020
aaaactcttt caggaggcat caaggtcaat gggcctcgtc tagagagcct agtgctgacc 1080
tatatcaatg ctatcagcag aggggatctg ccctgcatgg agaacgcagt cctggccttg 1140
gcccagatag agaactcagc cgcagtgcaa aaggctattg cccactatga ccagcagatg 1200
ggccagaagg tgcagctgcc cgcagaaacc ctccaggagc tgctggacct gcacagggtt 1260
agtgagaggg aggccactga agtctatatg aagaactctt tcaaggatgt ggaccatctg 1320
tttcaaaaga aattagcggc ccagctagac aaaaagcggg atgacttttg taaacagaat 1380

caagaagcat catcagatcg ttgctcagct ttacttcagg tcattttcag tcctctagaa 1440
gaagaagtga aggcgggaat ttattcgaaa ccaggggget attgtctctt tattcagaag 1500
ctacaagacc tggagaaaaa gtactatgag gaaccaagga aggggataca ggctgaagag 1560
attctgcaga catacttgaa atccaaggag tctgtgaccg atgcaattct acagacagac 1620
cagattctca cagaaaagga aaaggagatt gaagtggaat gtgtaaaagc tgaatctgca 1680
caggcttcag caaaaatggt ggaggaaatg caaataaagt atcagcagat gatggaagag 1740
aaagagaaga gttatcaaga acatgtgaaa caattgactg agaagatgga gagggagagg 1800
gcccagttgc tggaagagca agagaagacc ctcactagta aacttcaggt atccaaatgc 1860
aaaananaaa 1870

<210> 19
<211> 628
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:402231.6:2000MAY19

<220>
<221> unsure
<222> 580, 592
<223> a, t, c, g, or other

<400> 19
gcgctctctt ttccaggatc atccagcagc tcgtcaacgg catcatcacg cccgccacca 60
tccccagcct gggcccctgg ggagtcctgc actcaaaccc tatggactac gcctgggggg 120
ccaacggcct ggatgccatc atcacacagc tcctcaatca gtttgaaaac acaggccccc 180
caccggcaga taaagagaaa atccaggccc tccccaccgt ccccgtcact gaggagcacg 240
taggctccgg gctcgagtgc cctgtgtgca aggacgacta cgcgctgggt gagcgtgtgc 300
ggcagttgcc ctgcaaccac ctgttccaca caacatacga gcaggcctgg ctggagcagc 360
acgacagctg ccccgtctgc cgaaaaagcc tcacgggaca gaacacggcc acgaaccccc 420
ctggcctcac tggggtgagc ttctcctcct cgtcgtcatc gtcctcctcc agctcgccca 480
gcaacgagaa cgccacaagc aactcgtgag cccacgtcgg ccgtcgggaa agcacggggc 540
ctttcccacc caccctcagc cagcgccaca cggcacccan agactgggtg cnccggcggc 600
gccacgcttg gctggtcagc gctgcagg 628

<210> 20
<211> 798
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1076157.1:2000MAY19

<220>
<221> unsure
<222> 777
<223> a, t, c, g, or other

<400> 20
aaaaaaaaat tgctttatgg aagaaagtaa gtatagacag agagaaaggg atctgatgac 60
caaagcaggg aataaatgtt tggagtccac ggcatcctga gaacttcttg ggaatagagt 120
ctaggccccc aatgctgtca ctctcaccca tcctcctcta cacatgtgag atgtttcagg 180
acccagtggc ttttaaggat gtggctgtga acttcaccca ggaggagtgg gctttgctgg 240
atatttcgca gaggaaactc tacagggaag tgatgctgga aactttcagg aacctgacct 300
ctatagggaa aaagtggaaa gaccagaaca ttgaatatga gtaccaaaac cccaggagaa 360
acttcaggag tctcatagaa gggaatgtca atgaaattaa agaagacagt cattgtggag 420
aaacttttac ccaggttcca gatgacaggc tgaacttcca ggagaagaaa gcttctcctg 480
aagcaaaatc atgtgataac tttgtatgtg gagaagttgg cataggtaac tcatctttta 540
atatgaacat cagaggtgac attgggcaca aggcatacga gtatcaggac tatgcaccaa 600
agccatataa gtgtcaacaa cctaagaaag ccttcagata tcacccctcc tttagaacac 660
aagaaaggaa tcacaccgga gagaaaccct atgcttgtaa agaatgtgga aaaaccttta 720
tttcccattc aggcattcga agacgcatgg taatgcacag tggggatgga cccttanatg 780
taagttttgt gggaaagc 798

<210> 21
<211> 410
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1083142.1:2000MAY19

<220>
<221> unsure
<222> 51
<223> a, t, c, g, or other

<400> 21
ttccgttttc gcgtggttct tttgcaagct ctggattctc tggagtttga ntgtttccag 60
tattggaacc ccaccaagta ggactgatca ggtcttacaa ttctaaaacc atgacctgtt 120
ttcaggaatt agtgacattc agggatgtgg ccatagactt ctctcggcag gagtgggaat 180
acctggaccc taatcagagg gacttataca gggatgtgat gttggagaac tatagaaacc 240
tggtatcact gggaggacat tccatttcta aaccagttgt ggttgattta ctggagcgag 300
gaaaagagcc ctggatgatt ttgagggaag aaacacagtt cacagatttg gatttacagt 360
gtgagataat cagctacata gaagtaccca cttatgaaac agatatatcc 410

<210> 22
<211> 819
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1083264.1:2000MAY19

<400> 22
cggaagccga ttgcagggag aaactgtttt cgcagcagtg cgcctccctt ttccagccac 60
cggttctcct gaccccgagt gtgggggtg acttcagtct cctgacatcc agtgttctct 120
cgagccagtt tccagcccac agaaaatgag ctcttccgga agtgggcatc ttattccaat 180
cccctccctg tgaatgtgtg gagaaaaaga gatgggaacg aggcagagga aatagagaaa 240
ttttgaaaga gaaatgaaga atgagagacc cattaacaga aggcaaagta gaaggttcac 300
aaattttaag aaagggagaa taaagtgaaa aaaatctcag aaggaatcca ctcaacagac 360
gaggattcac ttccaaagag acatattatg caaggaagca acttggaaga ggaaagaaaa 420
gaagtcagga atggccctta ctcagggacc cttgaaattc atggatgtgg ccatagagtt 480
ctctcaggaa gagtggaaat gcctggaccc tgcgcagagg actttataca gggacgtgat 540
gttggagaat tataggaacc tggtctccct gggaatctgt cttcctgacc tgagtgttac 600
ctccatgtta gagcaaaaga gagatccctg gactctgcag agtgaagaga aaatagcaaa 660
cgatccagac ggcagggagt gcatacaaaa ggtgtgaaca cagagaggag ctctaaattg 720
ggaagtaatg caggaaacaa gaccttgtaa aaatcaaatt ggattcaact tttacagtat 780
aaattatgag tgatatacag ctaatttcaa gactgaaag 819

<210> 23
<211> 2516
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:350793.2:2000MAY19

<220>
<221> unsure
<222> 85, 118, 146
<223> a, t, c, g, or other

<400> 23
agtgtttctc atatctgggg tcactttaga caactgtgtt gaagttggac ggattgccaa 60
cacctacaat ctaaccgaag tgganaaata cgttaacagt ttcgtcttga agaatttncc 120
tgcattgctg agcacagggg agttcntgga aactcccttt tgagcgtctt gccttcgtgc 180
tttccagtaa tagccttaag cactgtactg aacttgagct ctttaaggct acctgtcgtt 240

ggcttcgcct ggaagagcct cggatggact ttgctgcaaa attaatgaag aacatacgat 300
ttccactgat gacaccacag gagctcatta attacgtgca aacggtggat ttcatgagaa 360
ctgacaatac ttgtgtgaat ttgcttttgg aagccagcaa ttaccaaatg atgccatata 420
tgcagccagt tatgcagtca gacaggactg ccattaggtc tgacaccact cacttggtta 480
cactaggagg agtgctgagg cagcagctgg ttgtcagtaa ggaattgcgc atgtatgatg 540
aaaaggccca tgagtggaaa tcgttagccc ccatggatgc cccaaggtac cagcatggca 600
tcgccgtcat tggaaatttt ctctatgtgg ttggcggaca gagtaattat gatacaaaag 660
gaaaaacggc agttgataca gtcttcagat ttgatcctcg atacaataaa tggatgcaag 720
ttgcatcttt aaatgaaaag cgcaccttct tccacctaag tgccctcaaa ggatatctgt 780
atgcagttgg tgggcgaaat gcagcaggtg aactgcccac agtagaatgt tacaatccaa 840
gaacaaatga atggacctat gttgccaaaa tgagtgagcc ccactatggc catgctggaa 900
ctgtgtatgg aggagtgatg tatatttcag gaggaattac tcatgatact ttccaaaagg 960
agctcatgtg ctttgaccct gatactgaca aatggatcca gaaggcgcca atgaccactg 1020
tcagaggtct gcattgcatg tgtacagtgg gagaaaggct ctatgtcatt ggtggcaatc 1080
acttcagagg aacaagtgat tatgatgatg tcctaagctg tgaatactat tcacctatcc 1140
ttgaccagtg gaccccaatt gctgccatgt taagagggca gagtgatgtt ggggtcgctg 1200
tcttcgaaaa taaaatctat gtggttgggg ggtattcttg gaataatcgt tgtatggtag 1260
agatagtgca gaaatatgat ccagataaag atgaatggca taaggttttt gatctgccag 1320
aatcccttgg tggcattcgt gcttgcacac tcacagtttt tccaccagaa gaaaccacac 1380
catcaccttc tagagagtcc cctctttctg caccttaaga tcatctctac aactaagatg 1440
ctgtagttct atctttgcaa tgtgtcataa attctcttct ttttcccccc taagtagtat 1500
atatgttagg attaccctct ggtaattgat acagatattg gaaaaaagac aacattgatg 1560
ttatttgtgc tctttgtttg gcctagaatg tttataaagt ggtaacacaa ccattctgga 1620
aatgtatccc atagaagctg atgtttaaca tatgaaaaaa aaagtattgt ctataaaatg 1680
tttcttcagt actttttaaa tgctgtgtat tgggtgtaag gtatttgtca tcttacatta 1740
gtaaacccaa taagccaagt tgaaggtgga ttatagtaaa tgtacaactg tgctcactag 1800
gcttcaagta aaaagttttc ctttcatctt tgactgtaag atgtcaaagg gaggcagcct 1860
gcttgaacag gaaacaatac acaaaaggtt gccaactcgc atgagctacc tccctctttt 1920
cataaagtat ttttgacata tctgtcaacc cacttgactg tgtgggtgca ttgagaacac 1980
aaagtttcct agacacacag gagaagtagc ttaaattcac taatattaat ttaaaaagca 2040
gcatgaaccc tctacttata aacaagggtt tggtgttttt aaagtgtgta tacatacata 2100
cacatacaca catgcacata tgtcaaatat aatttttta aaaattgagt ggcacatcaa 2160
agaaatgtga aattaaaaag aattcttcca aaaagcagct tccattaaaa tgggaattca 2220
gtatgcacat actgaatgca tatatgtaga accatacaga atttaggtgg ataagggcta 2280
gaaattttga gcaacaaaat ttgtcacttg accagatttt atcttcaaaa actgtattct 2340
actccttctc ctttgctgtt gaggtaactt gcatattata tgtattctgt atactcagtt 2400
cataaggtta tttagcacaa agtatagcag cttcacctgg agagctgctt ttgctcagta 2460
aattcaactt ccatgtttta tctttttttg ttccaataaa aacatttaat gtcaaa 2516

<210> 24
<211> 1660
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:408751.3:2000MAY19

<400> 24
tagggaccca ggatggcaga tccgggaccg ggctgggctg gcttggaaca tgcttgccaa 60
ctcagccagc gtgaggatcc tcatcaaggg aggcaaggtg gtgaacgatg actgcaccca 120
cgaggctgac gtctacatcg agaatggcat catccagcag gtgggccgcg agctcatgat 180
ccctggcggg gccaaggtga ttgatgccac aggaaaactg gtgatccctg gtggcatcga 240
caccagcacc cacttccacc agaccttcat gaatgccacg tgcgtggacg acttctacca 300
tgggaccaag gcagcactcg tcggaggcac caccatgatc atcggccacg tcctgcccga 360
caaggagacc tcccttgtgg acgcttatga gaagtgccga ggtctggccg accccaaggt 420
ctgctgtgat tacgccctcc acgtggggat cacctggtgg gcacccaagg tgaaagcaga 480
aatggagaca ctggtgaggg agaagggtgt caactcgttc cagatgttca tgacctacaa 540
ggacctgtac atgcttcgag acagtgagct gtaccaagtg ttgcacgctt gcaaggacat 600
tggggcaatc gcccgcgtcc atgctgaaaa tggggagctt gtggccgagg gtgctaagga 660
ggcactggat ttggggatca caggcccaga aggaatcgag atcagccgtc cagaggagct 720
ggaagctgaa gccactcatc gtgttatcac cagggatggg ggaaaccatg acgccgcctc 780
ctggtgcagt gcacaccatc tctatccctg tcagccctca ctgggtcatg ggccttgggc 840
agatgtcaaa gagcccagca gcagcggtgg tggccagctg ggcagagcat ccttgcttgg 900
gctaggaaag ctttaccttc tctgagtgcc tccgcctgag agatgtgtga cccgtggcac 960
cagggaacca cgtcttggag tggtccactg taggccatgc gcttcatcca cccccagtcc 1020
ctacataggc cctacccttg cccgggagct tctagataga aatcagaaag agattcaagg 1080

agccaaatga gcggtcagcc cccaccatgc actccttgcc ccgtgcagag ctccagccag 1140
cttcgtcacc agccccactg gctcctggtt ggaacgaaag ggtctctggt tgcactgaat 1200
gcagctctca aactggtctt gtacttgctg aataaatact gttgttcttg ccttagctgc 1260
tctctaggtt tgtggggtta agttgccaga aaattgtgct actgtgtgtg cgtgtgcgtg 1320
cgtgtgtgta gtgctaggag tccacagtag gtctctgtca agccgatgtc gtgatgaggg 1380
cttttctgat actgacccag aagccacaga accacaagga aacccaaacc ccctccagct 1440
gctgaggcgc aggcacagcc tggggtcgga tggagcctcc agcaccccag cacccaggtg 1500
acttccccac tcccctgtaa atgtcatggt gctaagactg tgtcaacccc aagacgacac 1560
atggtcctgt gctttggcca ccgtttgagg caaaaactaa acagcccgac acgttgtgtt 1620
ctggtgcagg tttgtattaa actgtagcta cttctcaaaa 1660

<210> 25
<211> 2762
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:336120.1:2000MAY01

<400> 25
gaagaccatg agggtgcaca gctggaaaac tctggtgtct cagcttaggg cctcctccgg 60
gaagagctaa ctgctcccag gtgaagccgg tgcccgcggg cggtccgtac accccgcagc 120
cggctcgcac cgctcgagag cctcggccgc tgtgtcttcc acgtctgcag ctcagccagg 180
gcgcgcaggg cgagtggggt ccactggcgg gtaaagggga ccaggacggc gaggatggac 240
gcacagacct ggcccgtggg ctttcgctgc ctcctccttc tggccctggt tgggtccgcc 300
cgcagcgagg gcgtgcagac ctgcgaagaa gttcggaaac ttttccagtg gcggctgctg 360
ggagctgtca gggggctgcc ggattcgccg cgggcaggac ctgatcttca ggtttgcata 420
tccaaaaagc ctacatgttg caccaggaag atggaggaga gatatcagat tgcggctcgc 480
caggatatgc agcagtttct tcaaacggtc cagctctaca ttaaagtttc taatatctcg 540
aaatgcggct gcttttcaag aaacccttga aactctcatc aaacaagcag aaaattacac 600
cagtatactt ttttgcagta cctacaggaa catggccttg gaggctgctg cttcggttca 660
ggagttcttc actgatgtgg ggctgtattt atttggtgcg gatgttaatc ctgaagaatt 720
tgtaaacaga ttttttgaca gtctttttcc tctggtctac aaccacctca ttaaccctgg 780
gtgtgactga cagttccctg ggaatactca gaatgcatcc ggatggctcg ccgggatgtg 840
agtccatttt tgtaaattat tccccaaagg agtaatgggg acagatgggg gaggtccctg 900
ctgcccagcc gcacttttct gcaggcactc aatctgggca ttgaagtcat caacaccaca 960
gactatctgc acttctccaa agagtgcagc agagccctcc tgaagatgca atactgcccg 1020
cactgccaag gcctggcgct cactaagcct tgtatgggat actgcctcaa tgtcatgcga 1080
ggctgcctgg cgcacatggc ggagcttaat ccacactggc atgcatatat ccggtcgttg 1140
gaagaactct cggatgcaat gcatggaaca tacgacattg gacacgtgct gctgaacttt 1200
cacttgcttg ttaatgatgc tgtgttacag gctcacctca atggacaaaa attattggaa 1260
caggtaaata ggatttgtgg ccgccctgta agaacaccca cacaaagccc ccgttgttct 1320
tttgatcaga gcaaagagaa gcatggaatg aagaccacca caaggaacag tgaagagacg 1380
cttgccaaca gaagaaaaga atttatcaac agcctttcga ctgtacaggt cattctatgg 1440
aggtctagct gatcagcttt gtgctaatga attagctgct gcagatggac ttccctgctg 1500
gaatggagaa gatatagtaa aaagttatac tcaagcgtgt ggttggaaat gggatcaaag 1560
cccagtctgg aaatcctgaa gtcaaagtca aaggaattga tcctgtgata aatcagatta 1620
ttgataaact gaagcatgtt gttcagttgt tacagggtag atcacccaaa gctgacaagt 1680
gggaacttct tcagctgggc agtggtggag gcatggttga acaagtcagt ggggactgtg 1740
atgatgaaga tggttgcggg ggatcaggaa gtggagaagt caagaggaca ctgaagatca 1800
cagactggat gccagatgat atgaacttca gtgatgtaaa gcaaatccat caaacagaca 1860
ctggcagtac tttagacaca acaggagcag gatgtgcagt ggcgactgaa tctatgacat 1920
tcactctgat aagtgtggtg atgttacttc ccgggatttg gtaactgaac tcttctgtcc 1980
tgacatacct tactgaagtc tcgatttctt ctctctctgc atatgcctgg aataagagat 2040
ccttttttcaa tgtaacaatt atatttatga aaagatatgt tacactaact tctcagaagc 2100
caagctgaaa tattcataaa gtccctaaaa ctcaacgttt aaatgacaca ctttaaaaat 2160
atgtcttttt tcaatctaac tgaaaacctt cttaacttct aatatattaa atctgaagat 2220
gtgaagggca cagaagtgac tttgaataag aagaatttag tgtatctgta attttattat 2280
caattcccaa gccccttcct ttctaaatta aaaatgtttt catttgaaag tgtatttgcc 2340
agacaatgaa aacagtatgc agtatttctt aaagtattga aattagaata tcatgaaata 2400
aatcaaaaca tacaatggca agtagtatgc atgcatattc aagagactct tccatttttg 2460
caagctgtag aaggaaatgt ctgaatgtct ataagttatg gggtagattc ttgagaagca 2520
tttccatata atttcactga agaaccttga taattttgac ccactgtaac ttagccactg 2580
atgaaccttа aagctgagta ttttattaac acctgatttg tattccatta tattcaaaat 2640
gcatctttgg tattgtgcct ctgctcccat ctctctcttt gcctcataga tttagctatg 2700
ttgggaagca catgcttgct ctaggaatat ctccaataaa gctgttaact atttggtgga 2760

aa 2762

<210> 26
<211> 4328
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:234104.2:2000MAY01

<400> 26
tgcgcccgga gccggggccg agtcgctgcc gcagctgttg gggcgcccgg gccaggcgac 60
gcggccgtcg cccgtgcccc tcccagaccg caccggccgc atggagcccc cggagggcgc 120
cggcaccgga gagatcgtta aggaggctga ggtgccgcag gctgcgctgg gcgtcccagc 180
ccaggggaca ggggacaatg gccacacgcc tgtggaggag gaggtcgggg gcatcccagt 240
accagcaccg gggctcctgc aggtcacgga gaggaggcag cctctgagca gcgtctcctc 300
tctggaggtc cacttcgacc tcctggacct cactgagctc accgacatgt cggaccagga 360
gctggccgag gtctttgctg actcggacga cgagaacctc aacaccgagt ccccagcagg 420
tctgcacccg ttgccccggg ccggctacct gcgctcccct tcctggacga ggaacaaggg 480
ctgagcagag ccacgagaag cagcccctag gcgaccccga gcggcaggcc acagtcctgg 540
acacgtttct cactgtggag aggccccagg aggactagac catctccacc tgccccagct 600
cctgcaggga tggggtccga acacgatggc agatctgggc cagtgctgac cccagcagac 660
acacttcacc cgcccacgag gctccagccg tcacctcctg acacacaccc tgggggcagc 720
tctctgccag ccccgagacc ggccttgtct tgctgggcac gggtcttcgc ctcacttgtg 780
agaccagccg gctttcctgg ggggacacac ggggcccccg gtatgcctct ggggagcccc 840
agcacaagca cagcccagtg gccttacgtc cagctcgttc ctgggccccg agtcaggaag 900
acagcgtcac ggagtcactg ccaggaacgt gctgaggaat ggagtggccc acggcggcct 960
tggggtgaag gggacccagg gcctgtgaca gccactccag gaactcctgg gggtgctcca 1020
acctccgcgt tttcctgtgc tgccaagctt cagaagccag atgcgggttt ggtagtggct 1080
aatgggacaa tgtgctgtcc agcaaagcac acatggagaa gcggccccaa aattcccatc 1140
cttgatttcc atcctgcccc ttcttctact ccacggagtg cgctgtctca ctagtggtcc 1200
ccctccacaa ggctcagcct ctaagacctg cacctgcttc tcttggcccc tgcgtgacag 1260
acaagtccat tccctcctta gctcagaaca ccaaatatca ccagactgcc taagagactt 1320
gatgacacct cccggaatgc tctcggggtt ggggttcacc tctccttgtc ctgcacccac 1380
tgctaggcca cattctcgtt tctgctcaca tcccattgcc cggctacaag gcctgcccac 1440
ggcccttaaa cttgctgggc aggtttggag ccccatggga ccccgtgggt ctctgtccag 1500
gagcagcaga ggaaggttga caggccctgc tccctctgct ctgggggtgt ctgggagccc 1560
cagctcacac cctcccaatg cttatatgct gaagctcaca gaatgggctt cttgcctgac 1620
agcgaagtca aagaatgagt ttaatatcaa agtgtaagct tactttccat ccccaagcca 1680
gactggcccc tgccccattt cccatgagca cacttctggg gaaggacaac aggctccctg 1740
gccttcactc tcagcagagc tttgggagat gcccccaggc atgcccgtga gctccttctg 1800
tgtacctgct cccacttcgt gagccacccg gctgcccctc cgcactgctg gcaaacccag 1860
ttccctgcct cagcccaggt ctccttcccc ggtttccagt cacacaagag cccagcagct 1920
ttctctttca gtcccataag gggcagcctt ttgtccctgg ccactcttat ctttccccac 1980
ttcattccac ccagtctccc tcccccgtcc ctgcccaaac gcgcgccccc ccgcccctcc 2040
cttggcccca gcgcccagcc ctgctctccg cgctcggcca gagggagcca gtccggagac 2100
ggccgcacct ggctggagag gctgggcggg cggatgggtg gaaactcgcg gacgcgggag 2160
ccgatctgga ccggagcagc cgcgagcaga atggagtctc ctaacagcct ctcggtgctg 2220
atgtgaaatt tgaccatctg attccagttt ttttcttttc cttttctttt ttgcatttcc 2280
ttccctcgcc atccgtcgtg tagtgaattg ttcagtcttg ctccgtttca agagaggaga 2340
tcatgattga gtgaagccac cccgtccgca gccaggaaaa gcacaaagaa gaaactgcaa 2400
caatggccaa gctgacagaa tccatgacta acgtcctgga gggcgactcc atggatcagg 2460
acgtcgaaag cccagtggcc attcaccagc caaagttgcc taagcaggcc agggatgacc 2520
tgccaagaca catcagccga gatcggacca aaaggaaaat ccagaggtac gtgaggaaag 2580
acggaaagtg caatgttcat cacggcaacg tgagggagac ctatcgctac ctgaccgata 2640
tcttcaccac attagtggac ctgaagtgga gattcaacct attgattttt gtcatggttt 2700
acacagtgac ctggctcttt tttggaatga tctggtggtt gatcgcatac atacggggag 2760
acatggacca catagaggac ccctcctgga ctccttgtgt taccaacctc aacgggttcg 2820
tctctgcttt tttattctca atagagacag aaaccaccat tggttatggc taccgggtca 2880
tcacagataa atgcccggag ggaattattc ttctcttaat ccaatactgt gttggggtcc 2940
attgtcaatg cttcactggt gggatagcct gtgtgtgaaa atctctcaaa cccaagaaaa 3000
gggcagagac cctggtcttt tccacccatg cagtgatctc catgcgggat gggaaactgt 3060
ggctgatagt tccgggtagg ggaccttagg aattcccaca ttgtggaggc ttccatcaga 3120
gccaagttga tcaaatccaa acagaactcg gaggggggagt tcatcccgtt gaaccagacg 3180
gatatcaacg tagggtatta cacgggggat gaccgtctgt ttctggtgtc accgctgatc 3240
attaggcatg aaattaacca acagagtcct ttctgggaga tctccaaagc ccagctgccc 3300

aaagaggaac tggaaattgt ggtcatccta gaaggaatgg tggaagccac agggatgaca 3360
tgccaagctc gaagctccta catcaccagt gagatcctgt ggggttaccg gttcacacct 3420
gtcctgaccc tggaggacgg gttctacgaa gttgactaca acagcttcca tgagacctat 3480
gagaccagca ccccatccct tagtgccaaa gagctggccg agttagccag cagggcagag 3540
ctgcccctga gttggtctgt atccagcaaa ctcaaccaac atgcagaact ggagactgaa 3600
gaggaagaaa agaacctcga agagcaaaca gaaagaaatg gtgatgtggc aaacctggag 3660
aatgaatcca aagtttagtg ccctagctgg gcaaaccctt ctcttctccc cccaacacaa 3720
tctttccttg tctctcattc tctttctttt tctgtctctc tggctttgtt ctttatttgt 3780
ttatatttaa tttttacatg accagaaaac aaatcttcaa ggtgtaaaat atctacctgc 3840
cctctctcag ttattcagat tgacaaggta gacatggatt tgatgaaagt gcaaagtgcc 3900
ctcatttgtg gcccaagcct ggtctcctcc caaaatacta cacatccaac tcctggagat 3960
ttcagttact tacctgcatg tgttgtacaa taccagatca ctcaaaaagg tgtgtcaaag 4020
attttacctg ggatatgaca agcaaggttt ctggtgccta tttattcatt cagtgagaca 4080
cagagtggag ccctcagttt tatggatccc aattcatttc atctactaca gggtgaggtg 4140
cttgccccca tgtgggtgtg gcagttacag ggcccaggtg agctgaagac aaaccactgt 4200
acatatatat gccttatgta attattttct ttttgtaatt agtaataaaa cccagcatgt 4260
acaaaagtac catagaacag aactgctaaa tactgtacat agatgtatca ttaatgtagg 4320
tttagata 4328

<210> 27
<211> 569
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:450887.1:2000MAY01

<400> 27
cgtcggttca cttctccagg aaagggttcg tactcatggc gccgccgcag ccaaagtcgg 60
gcctcttcgt tggcatcaac aagggtcatg tcgtcaccaa gcgcgagctg cctccccgcc 120
cgtgccaccg caaggggaaa tcaacgaaga gggtgtctat ggtcaggggc ctgatcagag 180
aggttgctgg gtttgctcct tatgagaagc gtatcactga gcttctgaag gttggcaagg 240
acaagcgtgc cctgaagctt gctaagagaa agcttggaac tcacaagagg gcaaagaaga 300
agagagagga gatggcgggc gtcctcagga agatgaggtc ggctggtacg cacactgaca 360
aaaagaaata gagagcattt caagttcatg gagctggctg ccagagatta tgttccagtg 420
tctgattttc catacatgta gaacctaata gacatgtcaa agtattatgt atcgaaccag 480
ctcatgggat tttgctcctt ccaatgcatc cagggtttat gtatcgaacc aatttatggg 540
atcttgctct tattctaatg catccatgg 569

<210> 28
<211> 3644
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:119992.3:2000MAY01

<220>
<221> unsure
<222> 2628
<223> a, t, c, g, or other

<400> 28
gacaatcttc aggacacact tgaagctgct agcttttta caaatattac ccgttttgga 60
tttctgtaaa gtatttctta tatcaggagt ctctttggat aactgtgttg aggttggacg 120
aattgctaac acctacaatc ttatagaagt ggataaatat gttaataatt tcatctctga 180
agaactttcc tgctttattg agtactgggg agtttctaaa actccctttt gaacgacttg 240
catttgtgct ttccagtaat agtcttaagc actgtaccga acttgaactc tttaaggcag 300
cctgtcgctg gctaaggttg gaagaccctc ggatggatta tgctgcaaag ttaatgaaga 360
atattcgatt tccactgatg acaccacagg atctcatcaa ttacgtgcag acagtagatt 420
tcatgagaac agacaatacc tgcgtgaatt tgcttttgga agctagcaat taccaaatga 480
tgccatatat gtcagccagt gatgcagtca gatagaactg gcaatcgaac tggattccac 540
tcacttggtt acattaggag gagttttgag gcagcagctg gttgtcagta aagaattacg 600
gatgtatgat gaaagggcac aagaatggag atctttagcc ccaatggatg ctccccgtta 660
ccagcatggt tattggctgt tcattggaaa ctttctttat gtagttggtg gtcagagtaa 720

ttatgataca aaaggaaaaa ctgctgttga tacagttttc agatttgatc ctcggtataa 780
taaatggatg caggttgcat cattaaatga aaagcgcaca ttctttcact tgagtgccct 840
caaaggacat ttgtatgcag ttggtgggcg cagtgcagct ggtgaactgg gcacagtaga 900
atgttacaac ccaagaatga atgagtggag ctatgttgca aaaatgagtg aaccccacta 960
tggtcatgct ggaacagtat atggaggctt aatgtatatt tcaggaggaa ttacccatga 1020
cactttccaa aatgagctca tgtgttttga cccagataca gataaatgga tgcaaaaggc 1080
tccaatgact acagtcagag gtctgcattg catgtgtaca cgttggagat aagctctatg 1140
tcattggtgg caatcacttc aagaggaaca agtgattatg atgatgttct aagctgtgaa 1200
tactattcac caacccttga ccagtggacc ccaattgccg ccatgttaag aggccaaaga 1260
tgatgttgga gttgccgtct tttggaaaat aaatttaatg ttgttgttga atattctggg 1320
aataatcgtt gtatggtaga aattgtccag aaatatgacc cagaaaaaga tgagtggcat 1380
aaagtttttg atcttccaga gtcacttggt gggcattcga gcctgtacac tcacagtttt 1440
tccacctgaa gaaaaccctg ggtcaccttc tagagaatca cctctttcag caccttcaga 1500
tcattcttag gtctaaggtg taacaccttt gcagtacgtc gatgggtgat ctaatacttc 1560
cccttcagtt gtatcttctt acagtgattg gtacagttat tagatataaa ggtaactgat 1620
gttattcgtc ttgtatggct tttagtatgt gctatcaagt ggctaacaaa tgcattctga 1680
aaatgtattt aacatagctg tgctaacaaa tgaaaaaaag acgtagaaaa atgtttagat 1740
gtctttttgt gatgttatat aaaattgtag atgactgtgg taaatgtgta attatgtcca 1800
ttatgcttca aagttgaagt tttcatcttt gactccaaaa tgtcagaggg aggccgctct 1860
aaactaaaaa taacgaaagg ttgccaagta ttaatactag ttacctccct cttttcgtag 1920
tttttgtcat gtctgtcaac ttactcgatt gtgtggttgc attcagaata tttgaagttt 1980
cttacgtaga cagaaataat aaaaatatta actaggaaaa aacagtatag caccaagcca 2040
gtatttggta tctctctcta gagcgagcaa gagagggaga gaggaggaaa aaatacacat 2100
aatacaaaca tacatgcatg cacacataca tacatatgta tacacacaca taatttgaaa 2160
actgattggc cacttcaacg atggctgaaa ttgtttttaa attgaagttt ctttcttcca 2220
caaagcagcc cgtttctatt caaatggaaa ttcagtacca gagaataaat gtctatgtag 2280
tcatactgaa tttagataga taagggctac aagcatacta aatcgagcaa ccaaatttgt 2340
catgtgacta aacccgttac ttcagatgaa gcttacatta ctgttttctg cttgtgtatt 2400
ttcccgtaga gtacttttac acagattggt aaaggttcag gtttcacgag aactgctttt 2460
gtgcagaaaa tttaggttct tttttccacc ttttttgggt cagtaaaact taatgaaaaa 2520
agcaaagaaa aaaaatattc tggaacaaag ctataagggt tttaaagttc agcctcccaa 2580
cgttaagtca tcctaacatg attattttgt gatttggggg tgcttgcncc tggtgctgtt 2640
ccagtccatg tggcatcctg agctgtgtga tctgcctcga ggctatgatc tgagcacgca 2700
ggagataaca ttttcttctg catcaagtga ggaaaaatgt gcttttgggc catgtctcaa 2760
agacaggacc aacttcagat ttcccaaaga agccagctac agagcctctg gaacactatg 2820
gtcttacaag cagtacttaa aatcaaccct cgagcctctt caatgccgaa aggtatcccc 2880
tatttggttg agaaccacat ggtaattttt aatgggactt tttatcagca aatggagtta 2940
caggaattct ctgtaatgag tgattctgaa gaggtacttt cctgggaata attatctacc 3000
tgaagaaaaa aaattttata tatacattgt gtgtgtgtgt aatacacaca cacacaagcc 3060
ccctaatacc tggaagattg tcagcatgta aatcaggaac aactttctcc cttattgaca 3120
atccccatta attaaaactc aggaaccaag gcaaaatgaa ttggcttcta gggggtctga 3180
accttactgc cccatacaag tgttgattca ttttaatgct gtttatgatt tctgcattgg 3240
cagaaaattt tcatactttc tatgtttttt ttaattactc agtttttttat tacctaaaaa 3300
taggcacatt tgagtacatt tgaaaagtag aaaaattaga aattattaac tttattgaat 3360
aagcaagaag tgcatcctaa tccctttgat tattaatgag gttgaatatt tgtgtgctat 3420
cggtagctgt gtttctttga tcaaatgttc ctgtcctttt gcccttctgt tatctgttgg 3480
gagttgcttt gtttttcgta tcaagttata gggatctctt tatataataa atgtaattta 3540
acttgcattt gcttgcattt atttcttccc tcaatctgtt gtagttttac aaaggcaacg 3600
ctgttcagtt aatttttgag atcaaatttg tctttttttt tttt 3644

<210> 29
<211> 2805
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:197241.2:2000MAY01

<220>
<221> unsure
<222> 325
<223> a, t, c, g, or other

<400> 29
ccccgttccc gattcctgta gtagcggctg tattgcagcc gcctgccgaa ctgacccggg 60
tctggggact ggcccctctg gcgccgttcg gtttctctta ttgccttcac tgaggatgag 120

tccctttgtg gctctatgtg gaccctgcgg aatccaccgg cgcagtttca tctagcgact 180
ggtcaccctt ggcaatttat ggatatttaa acagggtcag acagtgtgga cgggggagtt 240
ccccctcctc actccccctt ggtgcttgac tccaggaata atttataaac tgtggaattt 300
ttttaaactg aagaacttgt atttncgata tgaactttat agaagctatt tataactttt 360
tttggattta agctggccaa aaaattgcta taacagatat atacgtttta tactattgtc 420
aggcaggatt taacattatc ctaaaaaggt aatttattct ctgtaacttc ctcaatagca 480
cctttgtgtc ctggcttttt catttttaa aattagtttt cacgattctg aagtaagtgg 540
tataaaaaca gttagggatg agttcaccca tgcctgactg cacatcaaag tgtcgatccc 600
tgaagcatgc tttggaagtc ccttctgtgg taacaaaggg gagcgaaaac ccgattaagg 660
cccttctctc cacgtcattg ttacaaagct gccactatca aggatgtttt tggcaggaat 720
gccctccacc cctgtttcct cctcgtggag aagaaaggag tgttagattg gcttattcag 780
aaaggagtgg atctgttggt gaaagaccaa gagtctggat ggacagcctt gcaccagaag 840
cacttttta tggacatatt gattgtgttt ggtctctatt gaagcatggt gttagtctgt 900
atattcaaga taaagaaggc ttgtcagctt tggatcttgt aatgaaggat agaccaactc 960
atgtagtatt caagaatact gatcctacag atgtttatac ttggggcgat aatacaaatt 1020
ttaccctggg tcatggcaag ccagaatagc aaacatcatc cagagttggt ggatctgttc 1080
tccaggagtg ggatttatat caagcaggtg gtgctttgta aatttcactc cgtgtttctg 1140
tctcagaaag ggcaggttta tacctgtggt catggtcctg ggagggcgat tagggacatg 1200
ggagatgaac agacatgctt ggtccctcgg cttgtggaag gactgaatgg tcataattgt 1260
tcccaagtgg cagctgctaa ggatcatact gttgtattaa ctgaagatgg atgtgtttat 1320
acatttggtc taaacatttt tcatcaatta ggaattattc caccgccttc cagttgtaat 1380
gtacccagac agatacaggc aaaatatctg aaaggaagga caatcattgg cgttgcagca 1440
ggcaggtttc atacagtcct atggactaga gaagctgttt acactatggg actacatggt 1500
ggacaactcg gttgtttgct agatcccaat ggagaaaagt gtgtaactgc tcctcgtcag 1560
gtctctgccc ttcaccataa agacattgct ctgtctttgg ttgctgcaag tgatggagct 1620
acagtctgtg ttaccacaag gggagatatt tacttacttg cagactatca gtgcaagaag 1680
atggcttcta aacagttgaa cttgaaaaaa gttcttgtgt ctgggggtca tatggaatac 1740
aaggttgatc ctgaacattt gaaagaaaat gggggtcaaa aaatttgcat tcttgcaatg 1800
gatggagctg gaagggtgtt ttgctggaga tcagtcaaca gttctctgaa gcagtgtcga 1860
ttgggcctat ccacgtcagg gtcttcattt ctgatatggc tttaaataga aatgaaattc 1920
tatttgttaa cgcaaggatg gagaaggatt tagagggaga tggtttgaag agaaaagaaa 1980
gagttctgga aaagaaagag attttatcaa accttcacga ttcctcatca gatgtgtctt 2040
atgtctctga tataaatagt gtgtatgaaa gaattcgact tgagaaactt acctttgcac 2100
atagagcctg ttagtgtcag cacagatcca agtggatgca actttgcaat cctgcagtca 2160
gatcctaaaa caagccttta tgaaaattcc agctgtgtcc tcatcatcct tttttgaaga 2220
gtttggcaaa ctgttgaggg aagcagatga aatggacagc attcatgatg tgacatttca 2280
agttggcaat agactcttcc ctgcacataa atatattttg gcagtgcatt ctgatttttt 2340
ttcagaaatt gtttcttttc agatggtaat acttcagaat ttacagatat ttaccagaaa 2400
gatgaagatt ctgccagggt gccatctctt tgtggtagag aaggttcatc cctgacatgt 2460
ttgaatacct tttacaattt atatacacag atacttgtga ctttttaact ccatggcttc 2520
aaacccaaga atacacttaa acaaaaaccc agaagaacta tcagggaact ctgaattctc 2580
atttgaataa agtgaatttc catgaagatg ataaccagaa gtctgcattt gaagtttaca 2640
aaagtaatca agctcaaaca gttagtgaga ggcagaagag caaacctaaa tcttgtgaaa 2700
caaggcaaaa atattaggga agatgatcct gtaagaatgt tgcaaactgt gtggaagaaa 2760
ttcgacttca gtaatttgag tagtaggtta gatggagtca gattt 2805

<210> 30
<211> 572
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:406860.20:2000MAY01

<400> 30
gtttgtatgt gatgctggag atgactcggc cttcttcact gtcactgtca cagctggcac 60
tgttctcaag agctgtgctg ccagtgggga gggctgagga tctggcgggt gaggcaggag 120
aggcctgctg gccaagccta tgtgcccctc tccatgccca cccaccagcc ccaccagaga 180
ggattgtgca cccggcagcc cgctccctgg atctgcattt tggggctcca gggcgcgtgg 240
agctgcgctg tgaggtggcc ccagctgggt ctcaggtgcg ctggtacaag gacgggctgg 300
aagtggaggc atcagatgcc ctgcagctgg gtgccgaggg gcccacccgc accctgaccc 360
tgccccacgc ccagcctgag gacgccgggg agtatgtgtg tgagacccgg catgaggcca 420
tcaccttcaa tgtcatcctg gctgagcctc cagtgcagtt ccttgctcta gagacaactc 480
caagcccgct ctgtgttggc cccggggagc cagtggtgca ggagggcgag ggcctagagc 540
tccatgccga gggccccgcc gagtctctgc at 572

<210> 31
<211> 1082
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:142384.1:2000MAY01

<400> 31
ggcggacgtg ctgccgagta gtcccggaag cgaagcagcg atggcggaga gtccgactga 60
ggaggcggca acggcgggcg ccggggcggc gggccccggg gcgagcagcg ttgctggtgt 120
tgttggcgtt agcggcagcg gcggcgggtt cgggccgcct ttcctgccgg atgtgtgggc 180
ggcggcggcg gagtgtgggc ggggccgggg gcccggggag cggcctggct ccgctgcccg 240
ggctcccgcc ctcagccgct gcccacgggg ccgcgctgct tagccactgg gaccccacgc 300
tcagctccga ctgggacggc gagcgcaccg cgccgcagtg tctactccgg atcaagcggg 360
atatcatgtc catttataag gagcctcctc caggaatgtt cgttgtacct gatactgttg 420
acatgactaa gattcatgca ttgatcacag gcccatttga cactccttat gaagggggtt 480
tcttcctgtt cgtgtttcgg tgtccgcccg actatcccat ccacccacct cgggtcaaac 540
tgatgacaac gggcaataac acagtgaggt ttaaccccaa cttctaccgc aatgggaaag 600
tctgcttgag tattctaggt acatggactg gacctgcctg gagcccagcc cagagcatct 660
cctcagtgct catctctatc cagtccctga tgactgagaa cccctatcac aatgagcccg 720
gctttgaaca ggagagacat ccaggagaca gcaaaaacta taatgaatgt atccggcacg 780
agaccatcag agttgcagtc tgtgacatga tggaaggaaa gtgtccctgt cctgaacccc 840
tacgaggggt gatggagaag tcctttctgg agtattacga cttctattag ggtggctgca 900
aagatcgcct gcaccttcaa ggccaaacta tgcaggaccc ttttggagag aagcggggcc 960
actttgacta ccagtccctc ttgatgcgcc tgggactgat acgtcagaaa gtgctggaga 1020
ggctccataa tgagaatgca gaaatggact ctgatagcag ttcatctggg acagagacag 1080
ac 1082

<210> 32
<211> 2497
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:895427.1:2000MAY01

<220>
<221> unsure
<222> 1938
<223> a, t, c, g, or other

<400> 32
tagcctgcac ctgtacggtc tcgggggggct gcggccagcg ccgggggcca cccccaggga 60
cctctgctgc ctactgcaag tggatgggga ggccagggcc cgaacagggc cactgccacg 120
ggggccggac ttcctgctgg ctggaccaca ccttccacct ggagctggag gccgccaggc 180
tcctgcgcgc cctggtgctt gcgtgggacc ctggcgtgag aaggcaccgg ccctgtgccc 240
agggcaccgt gctgctgccc acggtcttcc gagggtgcca ggcccaacag ctggccgtgc 300
gcctggagcc tcaggggctg ctgtatgcca agctgaccct gtcggagcag caggaagccc 360
ctgccacagc tgagccccgc gtctttgggc ttgcccctgc cactgctggt ggagcgggag 420
cggccccccg gccaggtgcc cctacatcat ccagaagtgc gttgggcaga tcgagcgccg 480
agggctgcgg gtagtgggac tgtaccgtct ttgtggctca gcggcagtga agaaagagct 540
tcgggatgcc tttgagcggg acagtgcagc ggtctgccta tctgaggacc tgtaccccga 600
tatcaatgtc atcactggca tcctcaagga ttatcttcga gagttgccca ccccactcat 660
cacccagccc ctgtataagg tggtactgga ggccatggca ccgggcaccc cccaaacaga 720
gttcccccca ccactgaggg cacccgaggg ctcctacagc tgcctgccag atgtggaaag 780
ggccacgctg acgcttctcc tggaccacct gcgcctcgtc tcctccttcc atgcctacaa 840
ccgcatgacc ccacagaact tggccgtgtg cttcgggcct gtgctgctgc cggcacgcca 900
ggcgcccaca aggcctcgtg cccgcagctc cggcccaggc cttgccagtg cagtggactt 960
caagcaccac atcgaggtgc tgcactacct gctgcagtct tggccagatc cccgcctgcc 1020
ccgacaatct ccagatgtcg cgccttactt gcgacccaaa cgacagccac ctctgcacct 1080
gccgctggca gaccccgaag tggtgactcg gccccgcggt cgaggaggcc ccgaaagccc 1140
cccgagcaac cgctacgccg gcgactggag cgtttgcggg cggggacttc ctgacctgtg 1200
ggcgggattt cctgtccggg ccagactacg accattgtga cgggcagtga cagcgaggac 1260
gaggacgagg aggtcggcga gccgagggtc accggtgact tcgaagacga cttcgatgcg 1320

ccctttcaacc cgcacctgaa tctgcaaaga cttcgacgcc ctcatcctgg gatctggaga 1380
gagagctctc caaagcaaaa tcaacgtgtg ccttctgagc ccagatgacg gcggtgggga 1440
ccccggttag taaggaccgg gcgcccagtg gctaaggcgg tgccctggtg accaaggacg 1500
agccagacct gttgctcagg ccgagctcct gggttgccag cgagttacca cggggaccag 1560
tcgcgtgtat ggcttgagac ttcattccca gtttccaggg cccggctatt tggacactag 1620
ttgccaagtc tggggcctgg ggatttcacg ggaccagcgg cttgtgaccc atctttcctg 1680
agcaccaagg gcttcccctt ttgttgccac aaacggtcgt cctcgcgctt gctagcgctg 1740
gcctctcttg cctccccttg gccggggcaa caccagttac tgtgagcatc accctgggtg 1800
tgggtgagtc acctctagta cggccctctt gctgctgcca accaaatcag tattagcttt 1860
gagcactgca ctgtttctcc ctcccttggg acggacacaa agactaggca tgaggcactc 1920
tttgtggggg gcagcccnct atccctgggt tccaagcatg ggacacaggg ggtagcctgg 1980
gggcttatag acggaacaca gcttgtttcc ccctccactt tccccgggga aaaccccacc 2040
caatggcctt ttagcagcca atggagataa cagagttctg gcccccttccc atccccatct 2100
ccttgccccc cccttgcccc ccccccccgaa aaaaatgtga gcacgttaaa cccctcccct 2160
ttggaggggg ccccctgaag cgtcaggcct gggggcagtt tggtacggga acatatttac 2220
ttgcctccca tgcatgtgct gtgtgtgtct gtgaggcacg ggtgtgcgtg gacacagtct 2280
gaaggcaagg catggtgagg gctctattca tgggaccaca gcaggaggga gcagtttgcc 2340
atgccccacc caccctggaa tcccccatat atggtgcctc agtgggcccc cgagttccag 2400
tgggagagtg acggttccct cctgtctccc tcttcttttc cgcacctccg atctttgtgg 2460
ataataaata aatatgcaca ggttctgaaa aaaaaaa 2497

<210> 33
<211> 2876
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:757439.1:2000MAY01

<220>
<221> unsure
<222> 1472, 1495, 2463
<223> a, t, c, g, or other

<400> 33
cggaagccgc ggtagcggag aagactggag ctccgaggag ctgcatctgc ggcaacctgt 60
gtgctgacgc tacgtgcctc ctggcttccg acgtagctcg cagctcccca gtctcactcc 120
attccttccc cacctggcgc gcacctgctc aagaccaggg tcctgccaag cgctaggagg 180
gcgcgtgcca ggggcgctag ggaactgcgg agcgcgcgcg ccatggggcc gccgcctggg 240
gccggggtct cctgccgcgg tggctgcggc ttttccagat tgctggcatg gtgcttcctg 300
ctggccctga gtccgcaggc acccggttcc cgggggggctg aagcagtgtg gaccgcgtac 360
ctcaacgtgt cctggcgggt tccgcacacg ggagtgaacc gtacggtgtg ggagctgagc 420
gaggagggcg tgtacggccc ggactcgccg ctggagcctg tggctggggt cctggtaccg 480
cccgacgggc ccggggcgct taacgcctgt aacccgcaca cgaatttcac ggtgcccacg 540
gtttggggaa gcaccgtgca agtctcttgg ttgggcctca tccaacgcgg cgggggctgc 600
accttcgcag acaagatcca tctggcttat gagagagggg cgtctggagc cgtcatcttt 660
aacttccccg ggacccgcaa tgaggtcatc cccatgtctc acccgggtgc agtagacatt 720
gttgcaatca tgattcggca atctgaaagg cacaaaaatt ctgcaatcta ttcaaagagg 780
catacaagtg acaatggtca tagaagtagg gaaaaaacat ggcccttggg tgaatcacta 840
ttcaattttt ttcgttttct gtgtcctttt ttattattac ggcgggcaac tgtgggctat 900
tttatctttt attctgctcg aaggctacgg aatgcaagag ctcaaagcag gaagcagagg 960
ccaattaaag gcagatgcta aaaaagctat tggaaggctt tcaactacgc acactgaaac 1020
aaggagacaa gggaaattgg ccctgatggg agatagttgt gctgtgtgca ttgaattgta 1080
taaaccaaat gatttggtac gcatcttaac gtgcaaccat attttccata agacatgtgt 1140
tgacccatgg ctgttagaac acaggacttg ccccatgtgc aaatgtgaca tactcaaagc 1200
tttgggaatt gaggtggatg ttgaagatgg atcagtgtcc tttacaagtt tccctgatat 1260
ccaatagaaa tatctaatag tgcctcctcc catgaagagg ataatcgcag cgagaccgca 1320
tcatctggat atgcttcagt acagggaaca gatgaaccgc cttctggagg aacacgtgca 1380
gtcaacaaaa tgaaagtcta cagctggtaa aaccatgaag caaattctgg tggcagtgga 1440
tgttattcct catgttgaca acccaaccct tnttttgga agaagactgg aaaanctcct 1500
aatcaagaga ctgctgttcg agaaattaaa tcttaaaatc tgtgtaaata gaaaactgtg 1560
aaccattaag taataacaga actgccaatc agggcctagt ttctattaat aaattggata 1620
aatttaataa aataagagtg atactgaaag tgctcagatg actaatatta tgctatagtt 1680
aaatggctta aaatatttaa cctgttaact tttttccaca aactcattat aatatttttc 1740
ataggcaagt ttcctctcag tagtgataac aacattttta gacattcaaa actgtcttca 1800
agaagtcacg tttttcatct tataacaatt ttcttataaa aacatgttgc ttcttaaaat 1860

gtggagtagc ctgtaatcac tttatttat gatagtatct taatgaaaaa tactacttct 1920
ttagcttggg ctacatgtgt cagggttttt ctccaggtgc ttatattgat ctggaattgt 1980
aatgtaaaaa gcaatgcaaa cttaggcgag tacttcttga aaatgtctat ttaagctgct 2040
ttaagttaat agaaaagatt aaagcaaaat attcattttt tacttttct tatttttaaa 2100
attaggctga atgtacttca tgtgatttgt caaccatagt ttatcagaga ttatgggact 2160
gaattgattg gtatattagt gacatcaact tgacactaga ttagacataa aattccttac 2220
aaaaatactg tgtaactatt tctcaaactt gtgggatttt tcaaaagctc agtatatgaa 2280
tcattcatac tgtttgaaat tgcgtaatga ccagagtaag taacactgaa tattgggcca 2340
ttgatcctcc gttccatgaa ttagtctacc agaaaaaaaa tggttctgta aaaattagtc 2400
ctgttggaaa atggtttttc caaacaatgt ttactttgaa aattgagttt atgtttgacc 2460
ctnaatgggc gtaaaattac attagaataa acgtaaaatt gctgtgccgt gtaactgata 2520
aattattgtg aaatgcatta ttcactggtg tattgaaaaa agaagaggga gggagaatta 2580
ccaggtgcca ttaataataa agatttgaag ctatcattcc accaatagtt aaatttagag 2640
actaatttaa aatatgcaca tttaatttgt acatctgtga tggcttattg tatatagaat 2700
atttgtatac aaatatatag cagaatttag gcaaaaaata aaacagacat gtatttttgt 2760
gtgctgaatg gatgaaacca attgcattct tgtacactga tatacaaatg ctgtaaatat 2820
gtccccattt ttattgattc tctttaaata taaaatgtaa ataaaatatt ccaata 2876

<210> 34
<211> 1288
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:1144066.1:2000MAY01

<220>
<221> unsure
<222> 1243
<223> a, t, c, g, or other

<400> 34
ggggtgcgac gccgagggcg ggggagcgcg cgccgctgct cccggaccgg gccgcgcacg 60
ccgcctcagg aaccatcact gttgctggga ggcgacctgt acaaatccta agcgaatttt 120
ttggagcatt ttcaccccgg aaactcgcca tccagaagtg tgcttcccgc acagctgcag 180
ccatggggtc tgaggaccac ggcgcccaga aacccagctg taaaatcatg acgtttcgcc 240
caaccatggg agaatttaaa gacttcaaca aatacgtggg ctacatagag tcgcagggag 300
cccaccgggc gggcctgggc aagatcatcc ccccgaagga gtggaagccg cggcagacgt 360
atgatgacat cgacgacgtg gtgatcccgg ggcccatcca gcaggtggtg acgggccagt 420
cgggcctctt cacgcagtac aatatccaga agaagggcat gacagtgggc gagtaccgcc 480
gcctgggcaa cagcgagaag tactgtaccc cgcgggacca ggactttgac gaccttgaac 540
gcaaatactg ggaaggaacg ctcaccttgt gtctccccga tctacggggc tgacatcagc 600
ggctcttggt atgatgacga cgtggcccag tggaacatcg ggagcctccg gaccatcctg 660
gacatggtgg agcgcgagtg cggcaccatc atcgagggcg tgaacacgcc ctacctgtac 720
ttcggcatgt ggaagaccac cttcgcctgg cacaccgagg acatggtacc tgtacagcat 780
caactacctg cactttgggg agcctaagtc ctggtgagtg tctacactgg ccctgccgcc 840
ggccggaccg agagcccctc gggagggagt caatcccggg tacacggctg ggcgccgtgg 900
caggggcccc accaggtgag gccgcaaagg tcggcctatg acggctggag atcttccgga 960
ccgcctgggg tcacccacca gctttggggt gggggatgtg cacccccaga gccgaagctc 1020
ccaggccccc agagcttgcg ctttgtaccc cggagtgccc cccattgagc tgtgagcggc 1080
cccaggtgtc cccatggcca ggagcgtggt cttgagcctc ctgagctgcc caggctgtgc 1140
tgcctcacag ccaagtggag acgttcctgg tgaagggaca ctgtccatgc tgcccagagg 1200
ggcctggcca ggatgaccct gcagccgctc cctcgcagtc tcngccctgg cacgtctggg 1260
ccaggcccta cagttaggag ggcagggc 1288

<210> 35
<211> 5271
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:243660.4:2000MAY01

<220>
<221> unsure

<222> 3667
<223> a, t, c, g, or other

<400> 35
tgaccctgag cggcccccctg gagccacatg ccctgagagc ccaggacccg gacccccac 60
acccttggg ggtggtggaa tctggtaagg gtccgcctcc caccacggag gaggaggcct 120
ccggcccccc aggagagccc cggctggaca gtgagacaga gagtgaccat gatgatggct 180
ttcttttcaa taaggtctcc tgagatccag ttgcctctac cgcccggaaa acgtcggacc 240
cagtccctca gtgccctacc caaggaacgg gactcatctt ctgagaagga tggacgcagc 300
cccaacaagc gggagaagga ccacatccgg cggcccatga atgccttcat gatcttcagc 360
aagcggcacc gggccctggt ccaccagcgt catcccaacc aggacaaccg gaccgtcagc 420
aagatcctgg gcgagtggtg gtatgccctg gggcccaagg agaagcagaa gtaccacgac 480
ctggccttcc aggtgaagga ggcccacttc aaggcccacc cagcattgga agtggtgcaa 540
caaggaccga aagaagtcca gctcagaggc caagcccacg gagcctgggg ctggcaggag 600
ggcacaagga gacgcgggag cggagcatgt cggagacggg cactgctgcg tgcccctggg 660
tgtgtcctcg tgagctcctg tccgttgcag cccagacact cctgagctca gacaccaagg 720
ctccggggag cagctcctgt ggggcagaac ggctacacac agttggggga cctggctcag 780
cccggccccg agctttctcc cacagcgggg tacacagcct ggacggcgga gaagtagaca 840
gtcaggcgct acaggaactg acgcagatgg tgtctggccc tgcatcgtac tctggcccaa 900
agccttctac ccagtatgga gctccaggac cctttgcagc ccctggtgag ggaggtgcct 960
tggcggccac tgggcggccc ccgctgctgc ccacccgagc ttctcgttct cagcgtgcgg 1020
ccagtgagga catgacgagt gatgaggagc gcatggtcat ctgtgaggag gaaggggatg 1080
atgatgtcat tgctgacgat ggcttcggcc ccactgacct tgatctcaag tgcaaggagc 1140
gggtgaccga cagcgagagt ggggacagct ctggggagga cccagagggc aacaagggct 1200
ttggtcggaa ggtgttttca cctgtgatcc gttcctcctt tacccactgc cgccccccac 1260
tggaccctga gcccccaggg cccccggatc ctcctgtagc ctttggcaaa ggctatggtt 1320
ccgccccatc ctcctctgcg tcctcgcctg cttcctcctc agcctcggca gccacctcct 1380
tctcactggg ctcaggaacc ttcaaggccc aggagtctgg tcagggcagc acagcgggcc 1440
ccctacggcc cccacccccct ggggctgggg gtccagcgac accttccaag gcaacccggt 1500
tcctcccaat ggatcctgcc accttccggc gcaagagacc cgaaagtgtg ggtggcctgg 1560
agccaccagg cccctcagtc atcgcggccc ctcccagcgg aggaggaaac atcctgcaga 1620
cactggtgct gcccccaaac aaggaggagc aagagggcgg cggagccaga gtgccctccg 1680
cccccgcccc atcactggcc tatggggccc cagcagctcc cctgtcccgt cctgccgcca 1740
ccatggtcac caatgtggtg cggcctgtca gcagcactcc tgtgcccatc gcctctaagc 1800
ccttccccac ctctggccgg gctgaggcgt ctccaaatga cacagcaggt gccaggactg 1860
aaatgggcac tgggtctcgg gtgcctgggg gctccccgct gggtgtcagc ttagtgtatt 1920
cggacaagaa gtcggcagca gccacctcac cagccccaca cttggtggct ggacccctgc 1980
tgggcactgt ggggaaggcg cctgccactg tcactaacct actggtgggc accccggggt 2040
atggggcccc tgcgcccccct gctgtccagt tcattgccca gggggcccct ggtggtggga 2100
ccactgcggg ctcaggagca ggtgctggga gtggccccaa tgggccagta cccctgggca 2160
tcctgcaacc aggtgccctg ggcaaggctg ggggaatcac ccaggtacag tacatcctgc 2220
ccacgctgcc ccagcagctt caggtggcac ctgccccagc accagccccct gggaccaagg 2280
cagcggctcc catgcggccc tgcacccacc accagcatcc gtttcaccct cccaccgggc 2340
acttccacca acggcaaagt cctggctgcc actgcaccca ctcctggcat ccccatcctg 2400
cagtctgtac cctccgcccc acccccaaa gcccagtcag tttctcccgt gcaggccccg 2460
cccccgggtg gctcagccca gctgctgcct gggaaggtcc tagtgcctct ggccgccccct 2520
agcatgtcat ttgcggggtg gaaggggcgg gacagccaca tgccgacatg gtgagcccag 2580
cccttctcag tacctgtgca aaatggtgcc cagtccccca gaaagatcat ccagctgacc 2640
ccggtgccct gtgagcacac ccagcggcct ggtgcctgcc cctgaggacc ccagacacac 2700
tccctggacc cacctgctca atctcagaag gtccttgttg acctcactcc accagaatca 2760
cctatgtgca gtcagcgggc gggcacgcgc tgcccctggg taccagccct gcgtccagcc 2820
aggctggaac agtcacctcg tacgggccca cgagctctgt agctctaggc ttcacctcgc 2880
tggggcccag cggccccgcc ttcgtgcagc ccctgctctc agcaggccaa gccccactgc 2940
tggctcccgg tcaggtgggc gtgtcacctg tgcccagtcc ccagctgccg cctgcctgtg 3000
cagcccccgg aggtcctgtc ataacagcat tttactctgg cagccctgca cccacctcct 3060
cagcacccct tggcccagcc atcccaggcc cccccaagcc tggtctacac tgtggccacc 3120
agcacaaccc cacctgcagc caccattctg cccaagggcc cgccagcccc tgccactgcc 3180
accccagccc cgactagccc tttccccagc gccacagcag gttccatgac ctacagctta 3240
gtggcccccа aggcccagcg gcccagcccg aaggcccccc agaaagtgaa ggcagccatc 3300
gccagcattc ccgtggggtc ctttgaggca ggtgcctctg ggcggcctgg ccctgcaccc 3360
cggcaggcgt ctggagcctg gcccagtccg agagccaact gccccagagt ctgagcttga 3420
gtgggcagcc cacaccacca gcccctccac ccctgccaga gacctggact cccacggccc 3480
ggagcagccc ccacacgtgg ccccgcacac tgcgtgaagg agcaggacca tgcggccaag 3540
ggccctgaga cccatggcca gcaaattccc cagctcatct tcagactggc gcgtccctgg 3600
gcagggcctg gagaatcgtg gggagcctcc cactcctccc agcccggccc cagctccagc 3660
tgtagcnccc tggtggcagc agcgagagca gcagtgggcg ggcagccggg gacaccccgt 3720
gagcgcaatg gaggctgtgc tggtactggc aagaaggtga aggtgcggcc cccgaccccct 3780

gaagaagacc tttgactctg tggacaacag ggtcctgtca gaagtggact tcgaagagcg 3840
ctttgctgag ttgcctgagt ttcggcctga ggaggttgct gcctccccca acctgcagtc 3900
tctgggccac ctcacacccc gggcccatcc atgggcatct ttaccgcaaa gaagaggaag 3960
aactccacgg acctggattc agcacccgag gaccccacct cgcccaagcg caagatgaga 4020
agacgctcca gctgcagctc ggagccctac acccccaaga gtgccaagtg cgagggggac 4080
atcttcacct ttgaccgtac aggtacagaa gccgaggacg tgcttgggga gctagagtat 4140
gacaaggtgc catacgtcct cacctgcggc agcatccctg gaccacgcgc cagggccctg 4200
gtcatgcagc tctttcagga ccatggcttc ttcccgtcag cccaggccac agccgccttc 4260
caggcccgct atgcagacat ctttccctcc aaggttgtgt ctgcagttga agattcgtga 4320
ggtgcgccag aagatcatgc aggctgccaa ctcccaccgg agcagccccc tggagctgag 4380
gctcctctcc ctgtaccgcc ccccactggc accgctgcat gcacctgccc ccactcccag 4440
ccccgcaggg ggccctgacc ccacctcacc cagctcggac tctggcacgg cccagtgctg 4500
ccccgcacac tgcctcacac ccccagagtc ggggcctgga cagcctggct gggagtgggg 4560
ctccccagcc cttcccccca acccccaggt ccctccaaag gttgccacag gcagggtgag 4620
ggacccctcg agaagatgcc aggacttata gtacccccctc aggacatgga cagtatgtgg 4680
gggcaggaag gttatctcct cccgggtaaa gccatttgcg tcctctccag tttggggcgg 4740
aatgaggcct gctcctcttg taaatacccc cttccctctg aagctccctc ccggtgctgg 4800
ggggcagctg ccggggagag ctgcaggggc aagtctccct cctccaagcc cctgtacata 4860
acctggcagc gtgtgacctt cagagctttt cactttatgc aaaaatggct cctgtgaggg 4920
ctgcaaggct ggagggtggt gcaggccttg ggcccacagg gagtgcgcct gtggaatagg 4980
ggggagtttc atgcacccct ttttttcccc agaggggggct ggactcaggg ttagtttgag 5040
gggtgggggc tccctgcact ttgccacaag gccacgggga gggttttctc cttcacccca 5100
ttctgccctc ccaacttggg ttgtactttc taaagaaggt gattcccccg tggcccttgg 5160
gcccccttccc caaggaacaa aacattgttg atcatggtgc aatatttctt actgataccg 5220
agaagccgca atgagcgaga ttaaagcctg tttacacaaa aaaaaaaaag g 5271

<210> 36
<211> 6070
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:334386.1:2000MAY01

<400> 36
ggctagacta ggacatacca aggtggttaa ttgtttgatt gggtgtggag caaatattaa 60
tcatactgat caagatggtt ggacagcatt aagatctgct gcttggggtg gccatactga 120
ggtagtttct gcactacttt atgctggcgt aaaagtggat tgtgcagatg ctgatagccg 180
aacagctttg agagcagcag catggggagg acacgaggat attgtactga atttgctaca 240
acatggcgct gaagtgaaca aagctgataa tgaaggtaga actgctttga tagcagcagc 300
atacatggga catagagaga ttgtggaaca cctactggac catggagcag aagtaaatca 360
tgaggatgtt gatggcagga ctgcactctc tgtagctgca ctttgtgtgc ctgcaagtaa 420
agggcacgca tcagttgtta gccttttaat tgatcgaggt gctgaagtag atcattgtga 480
taaagatggc atgactccac tgctggtagc tggctatgaa ggacatgttg actgtggttg 540
acttgcttct agaaggggga gcagatgtag atcacacaga taacaatggc cgtacacccc 600
tcttagcagc agcgtctatg ggtcatgcat cagttgtaaa tacacttttg ttttggggtg 660
cagctgtgga tagtattgat agtgaaggta ggacagtcct cagtatagct tcagcacaag 720
gaaatgttga ggtggtacgt actctactgg atagagggtt agatgaaaat cacagagatg 780
atgctggatg gacacctttg cacatggcag cttttgaagg gcacagattg atatgtgaag 840
cacttattga acaaggtgct agaacaaatg agattgacaa tgatggacga atccctttca 900
tattagcttc acaagagggt cattatgatt gtgttcaaat attactggaa aacaaatcca 960
acattgatca aagaggttat gatggaagaa atgcactgcg ggttgctgca ttagaagggc 1020
acagggacat tgttgaattg ctttttagcc atggtgctga tgttaactgc aaagatgctg 1080
atggtcggcc tacactttat atcttggcct tagaaaatca gcttacaatg gccgaatatt 1140
ttttagaaaa tggtgcaaac gtagaagcaa gtgatgctga aggaaggaca gcacttcatg 1200
tgtcttgttg gcaaggccat atgggaaatg gtgcaggtcc tgatagcata ccatgcagac 1260
gtcaatgctg cagacaatga aaagcgctct gctttgcagt ctgcagcctg gcagggccat 1320
gtaaaagtgg ttcagcttct gattgagcat ggtgctgtag ttgaccatac atgtaaccaa 1380
ggtgcaactg cactctgtat tgcagcccag gaagggcaca ttggttgtgg ttcaggtctt 1440
attagagcat ggtgctgatc caaaccatgc tgatcaattt ggacgcactg ctatgcgtgt 1500
tgcagccaaa aatggacatt ctcagataat taaattatta gaaaaatatg gtgcatctag 1560
tttgaatggc tgttccccat ctcctgttca cacaatggag caaaaacctc tacagtcatt 1620
gtcttcaaaa gtgcagtcat taacaattaa atcaaatagc tctggtagta ctggtggagg 1680
ggatatgcag ccttcgttac gtggtttacc ctaatgggcc tactcatgct tttagttctc 1740
cttcagaatc tccagattct acagttgacc ggcagaagtc atcactgtca aataattccc 1800
tgaaaagctc aaaaaattca tctttgagaa ctacttcatc tacagcaacg gctcaaacag 1860

tgccaattga tagctttcat aacttgtcat ttacagaaca aattcagcag cattcattgc 1920
cacgcagtag taagtcgaca gtcaattgtt tccccatctt ccacaacaca gtccttagga 1980
cagagtcata attcaccaag tagtgaattt gagtggagtc aagtaaagcc cagtttgaag 2040
tcaactaaag caagtaaagg ggggaaatca gaaaattctg ccaagtctgg atcagctggg 2100
aaaaaagcga aacaaagtaa ttcttcacag ccaaaggttt tagaatatga aatgactcag 2160
tttgatagga agaggaccta ttaggccaaa tccgggactg gctggcaccg ccttaaacaa 2220
atgccagcag aatctcaatg caaaattatg ataccttcag ctcagcagga aattggtcga 2280
tctcaacagc agtttcttat tcaccaacaa agtggtggaa cagaagaaga gaaatggaat 2340
aatgacaaat ccaaattatc atcttcagag caaccaggtt tttcttggta gggtttcagt 2400
cccacgaaca atgcaagata gagggcatca ggaagtgttg gagggatacc cttcctcaga 2460
gacagaatta agcctttaaa caagctctga agcttcagat tgaaggttct gaccctagct 2520
tcaactataa aaaggaaaca ccattataaa agtttcctat tctgtgaaac agaaggacat 2580
tgtgatggag tggttcttca gctactggat gggaaacata tgcctgttga tttgctgaaa 2640
aaacataaaa aatgaagaat gtgatcttct ggcagtacag ttaccttaat tactgtaatg 2700
tgcctaaata gtaaggctgc cttctcaatg taacctctg tgcttaaaaa atttcatttt 2760
gtgtgctttg tattcactac acaggaataa gcacttttta aaaatgcaga tacatactgc 2820
agttccctga taaaagctga aaagaaaatt tgagtatttt aagttaaaat tgtgataaaa 2880
aaatgtgcat gtgccataat caaatatata tgaaaaggca gtgttccttg tatttatttt 2940
tttttctttt tgtggcaaaa gaaacttaaa catactgttt cagtcacatt gcattgtagt 3000
gtatggcctg tttcttgtat cttgacaaga cgtagctcaa taaacacata tcttgcaacg 3060
tgttctatgt tcactcacac cttcagcatt ggataaaaat catttcctat ataaatatca 3120
cattgaaatg aaaaaatgat tgctcgcgca gtttacaaca actcatttta tagtacttta 3180
gggacgctgc tttaagtagt tccgatctgg actctcccag tagaattctt ctcatctctg 3240
gctaaacatg tcagaacaaa caaccaacca gtctgttggc agaacaaagt cctatttcat 3300
ccgcctggga tacaatttca tctttccatt caccttttgtc attccacctc ctaagaagac 3360
agacttatca ttcctgaggc atgaaaattc tcagggacaa agccatgcct cagtcacatg 3420
tgtgtgcaga gagaaatgca cctgtctatc taagggtaga tttttgatcc ctggaataat 3480
tcattgacta aactgacctc ttcctcctgg gctaaataaa ttaattttgc tggcttctct 3540
ctcagcggtt tctattttgt aaattgctgc atgaccaaaa tagccccact caaaatcaat 3600
tggattaatt ttaatggttt ggttggatga atattctgga tgaatataaa atgtgctgcc 3660
cttcacagat gacaccactc ccctgtcaat catagcacat gtgtactttt tattgttact 3720
taatagtgat ggatttgcac ttttctatcc tcatactctt tcctgttttc ttctttgtac 3780
aattgcatgc aggagggctg gatgccaggg ttaagagaga tattcatgac aaggaaggta 3840
aaattggttc aaatgagcat gtgtcccaca gccttagtct ccttactctt aaatcagtgg 3900
agctgtagct tagatgggtc gtttatatgt ctggagaagt tgtcataaca gtttagaagc 3960
caaggttgtg gatttgatct tagaatgggc ctgttaatct tatagaaccc agaaattctg 4020
ttcttttcat gtactgttga tgaggaatga ggaaagagaa tttggagatt cagcacacag 4080
accattgcta tttctagaac aaaatcttta gaatatgtcc taataacaaa ggtcagtaat 4140
gtcatcttca tatatgaagg acacatgtgg atacattgtt cgtggaaata gaaatgtata 4200
attatgataa atgtttcact tgaatcttat tggtaaggct tttcttgctt ttatttttta 4260
aagtcagcaa aactcatttt gtctgtcatc tataagtcat agtgaggact acgagataag 4320
ttgataagtt caattgatat taaggtgaca tggcaatatt aataactcaa atgtgaatgt 4380
ttcaagtatt gaattcttgt ccctatggga catttattaa ataaaattaa tgggactctt 4440
acaaagtagc cttaaaagtg ttaatgagtc tattaataaa tatgaacaca tactattatt 4500
agaaccaact ttactcatat ctcaaataca gtacatttac attactggta aagggatgaa 4560
gctctaattt ctattacatg taatttтctt tagaaagaga accctgaaag ctgccagttt 4620
ttctattaac tttgaattat tgaaattgta tttatttaat tttattgttt ttacaaaatt 4680
gcaccttgtg tccaaagggc gaagatatga cattgcataa gggatttatg tttttcaaag 4740
agctaactgt tttcatatcc atactatata cacttgaagc aattggtagg aagagaacct 4800
attggaaaaa tacgattttc aaaagtaagt attcctcggg atgtttttat ataaattatg 4860
ttttggaata gaacataaat gactttgagt taggataagg atgataggat gggtgctgga 4920
gatgccctgc cttgcttatt tgtttctgtg ggggatttaa tactataaat gaaaatggct 4980
tctgctccat gtgggaaagt aaaaattgtg cctccaaata taaaactcct tcaatagata 5040
tttttgaaag ttaaaatctt taaattttat aaatcagctg ttgccacgta cacaactatg 5100
atggcatgtg cttataacat ttttataaca ttccgtctgt gtctattccc taaatatcat 5160
gagtttataa taacggaaag tgataaaaat acagtaatga aataaaaatc tggaatgttt 5220
ctaaggaaag gttttggaag gaaaaatgac ctgaaatcag tgcgtactgc aaaggtacag 5280
ccgaataatt tttgttcttt gttttcaccc ccaccctcac tggaattctc accaaaaata 5340
gtttgatacc ttaaaaaaca gaagtaaata cttttcctca atattgtttt gggctatgca 5400
aatatttgtt ttgcttaatg tcctccattt acacttgctc agcaaatggt agttgcaaac 5460
aaatgctttc tttttatttt tcccttgggt ttgagggtat gtaaatagcc aaaaatgtac 5520
attggaattt cacattgtaa aagttttatt ttatcccttg gtatgatatt actcaaaaaa 5580
tccctgtgta tatgaaagtg ccataataaa tatatttgct ttacagagaa gatcttgttt 5640
taatttttgtc cttgaaccag tagaacatgc atgatatgca tatagcataa acaactgtta 5700
gttgttttaa gttattatct taaataaatc ctgagcaaat gaatttggaa acattttgca 5760
aagaagaaag tgaaatataa cactgtccaa aggaaggtag aaaaacaaag atttactgtt 5820
tatgtcttcc taagcctttt taaagactta atgttctttt ccccccccg tgactgatta 5880

tatactatat cacaccatgt caggttttgt gcctctgaga attgcagtaa tccaataact 5940
tttgtatatg tgtgctcctt gatcatcaga atattatggc catcttatgg cggatatttt 6000
gggagtttat tgcaaacatg gtcattcatt ttctaaataa aatttgtgtg tttcttcact 6060
cagtaaaaaa 6070

<210> 37
<211> 3474
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:347572.1:2000MAY01

<400> 37
gtcattcagt ggatgtgatc tgtggctcac aggggacgat gtcaagctcc ttcctggctc 60
cttctcagcc ttgttgcctg taactggctg ctcagtccac cattgaggaa caggccaaga 120
catttttgga caagtttaac cacgaagccg aagacctgtt ctatcaaagt tcacgttgct 180
tccttggaat tataacacca atattactga agagaatgtc caacaacatg caataagttg 240
ctggcgagac aaatgtgtct agccctttt acaaggaaca gtccacactt gcccaagatg 300
tatccactac aagcaaactt cacgacatct ccacatgtca acgcttcagc tgtgcacggc 360
ttcttcaagc cataaaactg tgagtcttca ggttggtcat cacgaagcac agagagcaaa 420
ccggttgaac acaatttcta atatacaaat ggagccacca atcctaacag taactggaaa 480
acgtcgtaac ccagataatc cacaagaatg cttattactt gaaccaggtt tgaatgaaat 540
aatggcaaac agtttagact acaatgagag gctctgggct tgggaaagct ggagatctga 600
ggtcggcaag cagctgaggc cattatatga agagtatgtg gtcttgaaaa atgagatggc 660
aagagcaaat cattatgagg acttattggg gattattgga gaggagacta tgaagtaaat 720
ggggtaaata gtggatatga ttacagccgc ggccagttga ttgaagatgt ggaacatacc 780
tgttgaagag attaaaccat tgataggaac atcttcagcc ctatgtgagg gccaagttga 840
tgaatgccta tccttcctat atcagtccaa ttggatgcct ccctgctcat ttgcttggtg 900
atatgtgcgg gtagattttg gacaaatctg tactctttga cagttccctt tggacagaaa 960
ccaaacatag atgttactga tgcaatggtg gaccaggcct gggatgcaca gagaatattc 1020
aaggagtccg cagaacttct ttgtatctgt tggtcttcct tatatgactc taggattctg 1080
cggcaaattc catgctatac ggacccagga aatgttcaga aagcactctg ccatccccac 1140
agcttgggac ctggggaagg gcgacttcag agatccttat gtgcacaaag ggtaacaatg 1200
gacgacttcc tgacagctca tcatgagatg gggcatatcc agtatgatat ggcatatgcc 1260
ggccaacctt tttctgctaa ggaaatggag cttaatgaag gattccatga agctgttggg 1320
gaaatcatgt cactttctgc agccacacct aagcatttaa aatccattgg tcttctgtca 1380
cccgagtttt caacgaacga caatgaaaca gaaataaact tcctgctcaa acaagcactc 1440
acgattgttg ggactctgcc atttacttac atgttagaga agtggaggtg gatggtcttt 1500
aaacggggaa attcccaaag accagtgggt gaaaaaggtg gtgggagatg aagcgaaaga 1560
atagttgggg tgtgtggaac ctgtgcccca tgatgaaaca tatctgtgac cccgcatctc 1620
tgttccatgt ttctaatgat tactcattca ttcgatatta cacaaggacc ctgttaccaa 1680
ttccagtttc aaagaagcac ttttgtcaag cagctaaaca tgaaggccct ctgcacaaat 1740
tgtgacattc tcaaattcta cagaacgtcg tggacagaac actgttcaat atgctgaggc 1800
ttggaaaact cagaaccctg gaccctagca ttggaaaatg ttgtaaggac caaagaacat 1860
gaatgtaagg ccacctgctc aactactttg agcccttatt tacctggctg aaagaccaga 1920
acaagaattc ttttgtggga tggagtaccg actggagtcc atatgcagac cacagcatca 1980
caagtgagga taagcctaaa atcagctctt ggcagataaa gcatatgaat ggaacgacca 2040
atgaaatgta cctgttccga tcatctggtt ggatattgtt aattgaggca gtacttttta 2100
acaagtaaaa aatcagatga ttctttttgg ggaggaggat gtgcgagtgg ctaatttgaa 2160
accaagaatc tcctttaatt tctttgtcac tgcacctaaa aatgtgtctg gatatcattc 2220
ctagaaactg aagttgaaaa ggccatcagg atgtcccgga gccgtactcc atgatgcttt 2280
ccgtctgaat gacgacagcc tagagtttct ggggatacac ccaacacttg gacctcctaa 2340
ccagcccccc gtttccatat ggctgattgt ttttggagtt gtgatgggag tgataattgt 2400
tggccatggt catcctggat cttcactgga atcagagatc ggaagaagaa aaataaagca 2460
agaagtggag aataatcctt tatgcctcca tcgatattag ctaaggagta taaataatcc 2520
aggattccga aacactgatg atgttcagac ctccttttag aaaaatctat gtttttcctc 2580
ttgaggtgat tttgttgtat gtaaatgtta atttcatggt atagaaaata taagatgata 2640
aagatatcat taaatgtcaa aactatgact ctgttcagaa aaaatattgt ccaaagacaa 2700
caagtgccaa ggagagagca tcttcattga cattgctttc aagtatttat ttctgtctct 2760
ggatttgact tctgttctgt ttcttaataa ggattttgta ttagagtata ttagggaaag 2820
tgtgtatttg gtctcacagg ctgttcaggg ataatctaca atgtaaatgt ctgtctgaat 2880
ttcttgaagt tgaaaatcaa ggatatatca ttggagcata gtgttggatc ttgtatggaa 2940
tatggatgga tcacttgtaa ggatcagtgc ctgggaactg gtgtagcttg caaggattga 3000
gaatggcagt gcattagctc acttgtcact ggcatccatt ggtcaaggac tgacatgctt 3060
tccttcacag tgaactcagt tcaagtacta tggtgatttg cctacagtga tgttgtggaa 3120

tctgatctat gctttccttc aaggttgaca ggtcctaaag agagacagaa tccagggtac 3180
aggtagagga catttgcttt ttcacttcca aggtgtcttg tatcaacatc ttcctgtaca 3240
aacactgaaa tctagagctc aggggctctc gcgtgaatct cccagagaca tgcctgtata 3300
gaaatctcta tttctagctg ttctctaact gtcggagtga tatggaatat tccaactgta 3360
tgttcaccct ctgaagtggg tacccagtct cttaaatctt ttgtatttgc tcacagtgtt 3420
tgagcagtgc tgagcccaaa gcagacactc aataaatgct agatttaccc cctc 3474

<210> 38
<211> 3474
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:817314.1:2000MAY01

<400> 38
gctttcagag catcctcact ccgcccagtt cggtgccagc tgcgtgggct ccagcttcga 60
tcgttttcct tggaatgctc caaaactcag cagcgactaa gggaattcca ttggaatttg 120
ccgggcgtgc tctcaccccg cacggcaccc gcgccgtcag tcctcggatc ccatcacttc 180
agcccgaaga ttgcaacttt gcagagacga agaaatagca tggcatgaaa tatggctcag 240
ttctattaca aaagaaatgt taatgctccc tatagagacc gcatccctct aaggatagta 300
agagcagaat cagaactctc gccatcagaa aaagcctact tgaatgctgt ggaaaaggga 360
gattatgcca gtgtcaagaa atccctagag gaagctgaaa tttattttaa aatcaatatt 420
aattgcattg atcctctcgg aagaactgct ctcctcattg caattgaaaa tgagaacttg 480
gagctcatcg aactactctt aagctttaat gtctatgttg gagatgctct attacatgct 540
atcagaaaag aagtcgtcgg agctgttgag ctgttattga accacaaaaa acctagtgga 600
gaaaaacagg tgcctcctat actccttgat aagcagttct ctgaattcac tccagacatt 660
acaccaatca ttttggcagc ccatacaaat aattatgaga taataaaact cttggttcag 720
aaaggagtct cagtgcctcg accccacgag gtccgctgta actgtgtgga atgcgtgtcc 780
agttcagatg tggacagcct ccgtcactca cgctccagac tcaacatcta caaggccttg 840
gccagtccct ctctcattgc actgtcaagc gaagatcctt ttctcacagc ctttcagtta 900
agttgggaac ttcaggaact gagcaaggtg gaaaatgaat tcaagtcgga gtatgaagag 960
ctgtcacggc agtgcaaaca atttgctaag gacctactgg atcagacgag aagttccaga 1020
gaactggaaa tcattcttaa ttaccgagat gacaatagtc tcatagaaga acaaagtgga 1080
aatgatcttg caagactaaa attggccatt aagtaccgtc aaaaagagtt tgttgcccag 1140
cccaattgtc aacagctgct ggcatctcgc tggtacgatg agtttccagg ctggaggaga 1200
agacactggg cagtgaagat ggtgacatgt ttcataatag gacttctttt tcctgtcttc 1260
tctgtgtgct acctgatagc tcccaaaagc ccacttggac tgttcatcag gaagccattt 1320
atcaagttta tctgccacac agcctcctat ttgacttttt tgttcctgct gctgcttgcc 1380
tctcagcaca tcgacaggtc agacttgaac aggcaaggtc caccaccaac catcgtcgag 1440
tggatgatat taccgtgggt cctgggcttc atatggggag aaattaaaca gatgtgggat 1500
ggcggacttc aggactacat ccatgattgg tggaatctaa tggactttgt aatgaactcc 1560
ttatatttag caacaatctc cttgaaaatt gttgcatttg taaagtacag tgcccttaat 1620
ccacgagaat catgggacat gtggcatccc actctggtgg cagaggcttt atttgctatt 1680
gcaaacatct tcagttctct gcgtctgatc tcactgttta ctgcaaattc tcacctggga 1740
cctctgcaaa tatctctggg aagaatgctc ctggacattt tgaagtttct attcatatac 1800
tgccttgtgt tgctagcatt tgcaaatggc ctaaatcaat tgtacttcta ttatgaagaa 1860
acgaaagggt taacctgcaa aggcataaga tgtgaaaagc agaataatgc attttcaacg 1920
ttatttgaga cactgcagtc cctgttttgg tcaatatttg ggctcatcaa tttatatgtg 1980
accaatgtca aagcacagca tgaatttact gagtttgttg gtgccaccct gtttggggac 2040
attacaatgt catctctctg gttgttctac tcaacatgtt aatagctatg atgaataatt 2100
cttaccaact gattgctgac catgcagata tagaatggaa atttgcacga acaaagcttt 2160
ggatgagtta ttttgaagaa ggaggtactc tgcctactcc cttcaatgtc atcccgagcc 2220
ccaagtctct ctggtacctg atcaaatgga tctggacaca cttgtgcaag aaaaagatga 2280
gaagaaagcc agaaagtttt ggaacaatag gggtaagaac acagcatagg cgagctgctg 2340
ataacttgag aagacatcac caataccaag aagttatgag gaacctggtg aagcgatacg 2400
ttgctgcaat gattagagat gctaaagact gaagaaggcc tgaccgaaga gaactttaag 2460
gaactaaagc aagacatttc tagtttccgc tttgaagtcc tgggattact aagaggaagc 2520
aaactttcca caatacaatc tgcgaatgcc tcgaaggagt cttcaaattc ggcagactca 2580
gatgaaaaga gtgatagcga aggtaatagc aaggacaaga aaaagaattt cagcctttt 2640
gatttaacca ccctgattca tccgagatca gcagcaattg cctctgaaag acataacata 2700
agcaatggct ctgccctggt ggttcaggag ccgcccaggg agaagcagag aaaagtgaat 2760
tttgtgaccg atatcaaaaa ctttgggtta tttcatagac gatcaaaaca aaatgctgct 2820
gagcaaaatg caaaccaaat cttctctgtt tcagaagaag ttgctcgtca acaggctgca 2880
ggaccacttg agagaaatat tcaatctgga atctcgagga ttagcttcat cggggtgacc 2940
tgagcattcc cggtctcagt gaacaatgtg tgttagtaga ccatagagta aggaatacgg 3000

acacactggg gttacaggta ggaaagagag tcgtgtccat tcaagtcaga gaaggtgtgt 3060
ggtggaggac acggttccta taataccaaa ggagaaacat gcaaaagaag aggactctag 3120
tatagatcta tgatctaaac ctcccagaca cagtcaccca cgaagattac gtgaccacaa 3180
gattgtgata cttgaaggag gaagcgttta ccatacacat acgtattttc cgtagtgctc 3240
tgggtggggg aaaatgttta aattgtatta gcaaatgact aaattacact ttatagcgtt 3300
taatcaagat gtggaatatt acctgtaaca tgtttaaatt aaggcaaagg caatcaaaaa 3360
ccttttgtt ttgtagcctg cttttgcttt cacaatttgt cttacaattg tttttgttaa 3420
taaataaatg caccttgtat tcttgtactg ttgcaataac ccacagaaac attt 3474

<210> 39
<211> 1613
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:000290.1:2000MAY01

<400> 39
gtgagggctc ttgggttagt tcctgttagg ccccggccgg gggagtaggt tgaagtctcc 60
taacgatgcc cggtgggctg cgggcaccgg gagctgtgaa gggaacgtga gggggcggcg 120
tagtggagac ccacggcagg cctgaagaag agcggcggcc gagcccgcct tccctgcacc 180
atgctcatag aggatgtgga tgccctcaag tcctggctgg ccaagttact ggagccgata 240
tgtgatgctg atccttcagc cttagccaac tatgttgtag cactggtcaa gaaggacaaa 300
cctgagaaag aattaaaagc cttttgtgct gatcaacttg atgtcttttt acaaaaagaa 360
acttcaggtt tcgtggacaa actatttgac agtctctata ctaagaacta ccttccactt 420
ttggaaccag taaaagcctg agccaaaacc actagttcca agaaaaagac gaaattaccg 480
aagaggtatt tcaggagcca gcagaggaag aacgagatgg cagaaaaaag aaatatccta 540
gtccccagaa gactcgttca caatctactg aacgaaggac acgtgagaac aaaagagacg 600
acgggacctt ggcgagacta tgaccggtac tatgagcgga atgaattgta ccgtgagaag 660
tatgactgga gaagaggcag gagtcagagt cggagtaaga gtcgaggcct gagtcgcagt 720
agaagccgaa gtaggggcg cacgcaaaga ccgggatcca aataggaatg ttgagcacag 780
ggaaagatcg aagtttaaga gtgacaggaa tgacctggag aagttcctat gtgcctgtgt 840
ctgcaccacc tccaaactct tctgagcagt attcctctgg ggcacagtct attcccagca 900
ctgttactgt gatcgcaccc tggtcaccca ctcttgaaaa cacaacttgg gagttggtct 960
tattactatt aaaaattatt agctcttcca attcttttgg tcgaaaccta ccaccaaaga 1020
ggcgatgcag agattatgat gaaagaggat tttgtgtact tggtgacctt tgtcagtttg 1080
atcatggaaa tgatccccta gttgttgatg aagttgctct gccaagtatg attcctttcc 1140
caccccctcc tcctgggctt cctcctccaa caactcctgg aatgttaatg cctccaatgc 1200
caggtccagg cccaggcccg ggcccaggtc caggcccagg cccgggccca ggtccaggtc 1260
ctggccatag tatgagactt cctgttcccc aaggacatgg tcagcctcca ccatccgttg 1320
tgcttcccat accaagacca cctataacac aatcaagctt gataaacagc cgtgaccagc 1380
ctgggacaag tgcagtgccc aatcttgcat cagtgggaac aagactacct cctcctttac 1440
cccagaacct cctttacaca gtatcagaac gacagcccat gtactctcgt gaacatggtg 1500
ctgctgcatc tgagcgactt cagttgggga caccgcctcc tctgttggca gctcgtttgg 1560
tgccacctcg aaacctcatg ggatcctcca ttggatacca tacctcagtc tcc 1613

<210> 40
<211> 1056
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:023518.3:2000MAY01

<400> 40
ccagaggaaa ctagtcacaa aaaccctgac tatcacctga tagattgctt gtgctgcctg 60
ataattactc gcacttttcc caggctagtg caaatcttca ggggccgtcc aggactacag 120
agctgtttca ccctaccttg gcttcaatct cttcccccat gctcgaaggt gcggagctgt 180
acttcaacgt ggaccatggc tacctggagg gcctggttcg aggatgcaag gccagcctcc 240
tgacccagca agactatatc aacctggtcc agtgtgagac cctagaagct ccattcttcc 300
aagactgcat gtctgaaaat gctctagatg aactgaatat tgaattgcta cgcaataaac 360
tatacaagtc ttaccttgag gcattctata aattctgtaa gaatcatggt gatgtcacag 420
cagaagttat gtgtcccatt cttgagtttg aggccgacag acgtgctttt atcatcactc 480
ttaactcctt tggcactgaa ttgagcaaag aagaccgaga gaccctctat ccaacctttc 540
ggcaactcta tcctgagggg ctgcggctgt tggctcaggc ggaagacttt gaccagatga 600

agaacgtagc ggatcattac ggagtataca aacctttatt tgaagctgta ggtggcagtg 660
ggggaaagac attggaggac gtgttttacg agcgtgaggt acaaatgaat gtgctggcat 720
tcaacagaca gttccactac ggtgtgtttt atgcatatgt aaagctgaag gaacaggaaa 780
ttagaaatat tgtgtggata gcagaatgta tttcacagag gcatcgaact aaaatcaaca 840
gttacattcc aatttataa cccaagtaag gttctcaaat gtagaaaatt ataaatgtta 900
aaaggaagtt attgaagaaa ataaaagaaa ttatgttata ttatctagac tacacataag 960
taagccacac tatatcttca tgagttgcaa atccatggaa acacagtaaa ccaggcctga 1020
aacaaagcat ttccttggtt tcagtggtat tagatc 1056

<210> 41
<211> 3806
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:1084246.1:2000MAY01

<400> 41
cgttacaagc agtgcaggtt taccaacggc ttggggcagc gatatactaa acaaatttaa 60
tttaaaagca actgtgtgac gattcctcca agcaagaaat tggaattgaa tgtctcaagt 120
ctcgttgcgg ttgctgaggg gattggatat agggacctgg actccaacat gaagaagcta 180
gggagaattc atccaaacag gcaagtgttg gcctttattt tgatggtgtt cttgtctcag 240
gttcgcctcg agcctattcg ttattctgtg ttggaggaaa cagagagcgg ctcctttgta 300
gcccatctgg ccaaggatct gggcctggga attggggaac tggcctcccg gtcagcccgg 360
gtgctgtctg acgatgacaa gcagcgtttg cagctggatc gtcagactgg agatttgctt 420
ctgagggaga aactagaccg ggaagagctc tgtggtccta ttgaaccgtg tgtactgcat 480
ttccaagtgt tcctggaaat gccggtgcaa ttttttcata ggagaattat tgatccagga 540
tcatatatgt atcactctcc aatattccct gaaagggaag tgctcttgaa aatactagaa 600
aatagtccag ccgggtactc tatttccgtt gctaatagct gaggatttgg atgtgggcag 660
caatggtctt caaataatac acaatcagcc ccaattctca ttttcacatt ctcactcgaa 720
atcatagtga gggcaagaaa tacccagatt tggtgcagga caaaccacta gatcgagagg 780
agtcagcctg agttacagct taaccctcgt ggcgctggat ggtgggtcac cacctaggtc 840
tggcacggtc atggttcgaa tcctgatcat ggacatcaat gacaatgctc ctgagtttgt 900
gcacactcca tatggggtgc aggtcctgga aaacagcccc ctagactctc caattgttag 960
ggtcttagct agagatatag atgctggaaa cttcgggagt gtttcttatg gcttattcca 1020
agcatcagat gaaattaaac aaactttctc aataaatgaa gtcacgggag aaatactgtt 1080
gaaaaaaaaa ttggatttcg aaaaaattaa atcttaccat gtagaaattg aggccacaga 1140
tggaggaggc ctttctggaa aaggcactgt agtcatagag gtggtggatg tgaatgacaa 1200
tcccccagaa cttatcatat cttcactcac cagctccatc ccagaaaatg ctcctgagac 1260
ggtagtctct atcttccgaa ttcgagatag agattccgga gaaaatggaa agatgatttg 1320
ctctattcca gataatctac cgtttattct aaaaccaact ttgaagaatt tttacaccct 1380
ggtaacgggg gtgaccactg gaccgagaga ccagcactga gtacaacatc accatcgccg 1440
tcactgactt ggggacaccc aggctgaaaa cccagcagaa cataaccgtg caggtctccg 1500
acgtcaatga caacgccccc gccttcaccc aaacctccta caccctgttc gtccgcgaga 1560
acaacagccc cgccctgcac atcggcagtg tcagcgccac agacagagac tcgggcacca 1620
acgcccaggt cacctactcg ctgctgccgc cccaggaccc gcacctgccc ctcgcctccc 1680
tggtctccat caacgcagac aacggccacc tgttcgccct caggtcgctg gactacgagg 1740
ccctgcaggc gttcgagttc cgcgtgggcg cctcagaccg cggttctccg gctttgagca 1800
gcgaggcgct ggtgcgcgtg ctggtgctgg acaccaacga caactcgccc ttcgtgctgt 1860
acccgctgca gaatggctcc gcgccctgca ccgagctggt gccccgggcg gccgagccgg 1920
gctacctggt gaccaaggtg gtggcggtgg acggcgactc gggccagaac gcctggctgt 1980
cgtaccagct gctcaaggcc acggagcctg ggctgttcgg cgtgtgggcg cacaatggcg 2040
aggtgcgcac cgccaggctg ctgagcgagc gcgacgcagc caagcacagg ctcgtggtgc 2100
ttgtcaagga caatggcgag cctccgcgct cggccaccgc cacgctgcac gtgctcctgg 2160
tggatggctt ctcccagccc tacctgcctc tccctgaggc ggccccggcc caggcccagg 2220
ccgactctct caccgtctac ctggtggtgg cgttggcctc ggtgtcgtcg ctcttcctct 2280
tctcggtgct cctgttcgtg gcggtgcggc tgtgcaggag gagcagggcg gcctcggtgg 2340
gtcgctgctc ggtgcccgag ggccccttc cagggcatct ggtggacgta agcggcaccg 2400
ggaccctgtc ccaagagcta ccagtacgag gtgtgtctga caggagactc tgggactggt 2460
gagttcaagt tcctgaagcc aatatttcct aatctcttgg ttcaggacac cggggaggga 2520
agttaaggaa aacccccaag ttcagaaata gcttggtatt cagttaagta ttgtatttag 2580
ttcagtgaac cgcccgttaa gttttgtcaa acttcccact ggcaatgcct ttatttaaaa 2640
aaattgtcta cttatctgaa atattcatac cacaatttca aacctactca tgtccctgat 2700
aaagctaaat ttgtcccttt tttattgtta ttaattgcac ttaacatttt tagttatact 2760
ggatattgag tatggatttt ctctatattt gatctattgg tgattaatct ttttgtaatc 2820
ataaattact caattaggat aaaaataaat tatgttttaa tgaaattctt aaattaacat 2880

cttttaatg gaacatttaa gtgaatatat gaatattgaa tttctaaata tttgttgtgc 2940
ctgtctttac catgtaactt aatgtttgca aggccagagt gtttgaaagt tttgtattta 3000
actttataat taccttgtcc tttctggttg actatactag gctaagccct cttaatagcc 3060
atgagtgtaa aatttagttt actcattttt cacaaattgt aaattaacat ggcacttcac 3120
tacattggta atacactaaa attgtggtcc ttttcctctt gtgaccacca catgtctagt 3180
gattattttg tttatttggt tgctacttac ctagcacatt gtaatgttcc atgaatgcta 3240
atattaaatt ttgtaaaaat aacttattta taaataattt ttaaagagaa aaatctcata 3300
taatttgtca taacctttca ataaataaaa ctgttaaatc atgggcctga tatcatctta 3360
aaaaaaaatc ctcagaatct gaaataagcc ctaaatttct ccccaaaatc aagactcttg 3420
agagcatcat aggtctcctt gtgctacctt ttactcccta taaatagaaa tccaagtata 3480
ctttaatatg tgtatatttt ttggttttcc tacagcttct ccccatcttt caaaagaatc 3540
acgaaatttc ttctgcacct tggctattct gtttaaatct gataatcagt tgatctcagg 3600
tttttcactg tacattactt tgcagatatg gacagccttt acaaaaataa tttttaaatg 3660
cttaattatt ttaatttgtt ctttaaggta accttcagtt attttgtatt aatttaactt 3720
ctcaattatg ccaaagttgc acttgcatga aataaatatt attttgtcct tgtatagact 3780
ggaacagtaa taaatttatc tgaatt 3806

<210> 42
<211> 6230
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:1165828.1:2000MAY01

<220>
<221> unsure
<222> 4042, 4725
<223> a, t, c, g, or other

<400> 42
ctcgcttttc ttgcaatatt ttataccttt tcaattcata gaattactca agaaaactac 60
ctcagttggt tgctactttt tgttgattcc ttttaccaga catgactaag tttcttttc 120
atcagtagat ttctgggctc ctatattcac tagagattgc aactcctgga tttctcttac 180
actagaatcc tatttcgagc catatgggag attctgaatt ccagaacaaa agaattttgt 240
aatttaaaat tcgtgattgc tcaatggaat cattttaatt gttacttcat ttctgtcgtt 300
atttaaaact taagtggaga gttttctcag ggataagaaa accacaatca aggtcataca 360
aaacttttag aggcagtcag tctgctaaga aggctccagc aagagaaacg ggatcttctg 420
tttcaacaat cattacttaa gaaaaaatta agaaaatgaa ataagttttg cagaataact 480
gtgaaatttt tattcatgaa atatgtactt acactttggg ccacgtgatg tcactctttg 540
ccgcgatgtt ctctctgaat ccagacaaat acagcccttt tcccatggga aagaggctca 600
attctttttc actctctctg tgctgaacga tggcgaacac agcagaatgg gactgacgaa 660
atcagatgat ttcttctaat ttggaggcaa ttttcactaa ttagaagaag actgagtatt 720
tgaaatgtta tactcaagtc gaggagatcc agagggtcag cctctactgc tctcgcttct 780
gatcctcgca atgtgggtgg tggggagcgg ccagctccac tactccgtcc cggaggaagc 840
cgaacacggc accttcgtgg gccgcatcgc gcaggacctg gggctggagc tggcggagct 900
ggtgccgcgc ctgttccagt tggattccaa aggccgcggg gaccttctgg aggtaaatct 960
gcagaatggc attttgtttg tgaattctcg gatcgaccgc gaggagctgt gcgggcggag 1020
cgcggagtgc agcatccacc tggaggtgat cgtagacagg ccgctgcagg ttttccatgt 1080
ggacgtggag gtgaaggaca ttaacgacaa ccctccagtg ttcccagcga cacaaaagaa 1140
tctgttcatc gcggaatcca ggccgcttga ctctcggttt ccactagagg gcgcgtccga 1200
tgcagatatc ggggagaacg ccctgctcac ttacagactg agccccaatg agtatttctt 1260
cctggacgtg ccaaccagca accagcaggt aaaacctctt ggacttgtat tacggaaact 1320
tttagacaga gaagaaactc cggagcttca tttattgctc acggccaccg atggaggcaa 1380
acccgagctg actggcaccg ttcaattact catcacggta ctggacaaca atgacaatgc 1440
cccagtgttc gacagaaccc tgtatacggt gaaattacca gaaaacgttt ctatcggaac 1500
gctggtgatt caccccaatg cctcagattt agacgaaggc ttgaatgggg atattattta 1560
ctccttctcc agtgatgttt ctccagatat aaaatccaag ttccacatgg acccccttaag 1620
tggggcaatc acagtgatag gacatatgga ttttgaagaa agtagagcac acaagatccc 1680
agtcgaggct gtcgataaag gcttcccacc cctggctggt cattgtacac ttcttgtgga 1740
agttgtggat gtaaatgaca atgctccaca gttgactatc aaaacgctct cggttcctgt 1800
aaaagaggac gcacaactgg ggacagttat tgccctgatt agtgtgatcg acctagacgc 1860
agatgccaac gggcaggtga cctgctccct gacgccccac gtccccttca agctggtgtc 1920
cacctacaag aattactact cgttggtgct ggacagagct ctggaccgcg agagtgtgtc 1980
cgcctacgag ctggtggtta ccgcgcggga cgggggctcg ccttcactgt gggccacggc 2040
cagggtgtct gtggaggtgg ccgacgtgaa cgacaacgca ccagcgttcg cgcagtccga 2100

gtacacggtg ttcgtgaagg agaacaaccc gccgggctgc cacatcttca cggtgtctgc 2160
gcgggacgct gacgcgcagg agaacgccct ggtgtcctac tcgctggtgg agcggcggtt 2220
gggcgagcgc tcgctgtcga gctacgtgtc agtgcacgcg gagagcggca aggtgtacgc 2280
gctgcagccg ttggaccacg aggagctgga gctgctacag ttccaggtga gcgcgcgcga 2340
cgcgggcgtg ccgcctctgg gcagcaacgt gacgctgcag gtgttcgtgc tggacgagaa 2400
cgacaatgcg ccggcgctgc tgacacctcg gatgaggggc actgacggcg cagtgagcga 2460
gatggtgctg cggtcggtgg gcgccggcgt agtggtgggg aaggtgcgcg cagtggacgc 2520
cgactcgggc tacaacgcgt ggctttcata cgagctgcag ccagaaacgg ccagcgcgag 2580
catcccgttc cgcgtggggc tgtacacggg cgagatcagc acaacgcgtg ccctggacga 2640
aacggacgca ccgcgccagc gcctactggt gctggtgaaa gaccacgggg agccagcgct 2700
gacggccacg gccactgtgc tggtgtcgct ggtggagagc ggccaggcgc caaagtcatc 2760
gtcgcgggcg tcagtgggtg ccacgggccc cgaggtgacg ctggtggatg tcaacgtgta 2820
cctgatcatc gccatctgcg cggtgtctag cctgttggtt ctcacgctgc tgctgtacac 2880
tgtgctgcgg tgctcggcga tgcccaccga gggcgagtgc gcgcctggca aggccgacgc 2940
tggtgtgttc tagcgcggtg gggagttggt cgtactcgca gcagagggag gcagagggtg 3000
tgctctggcg agggtaagca gaagaccgac ctcatggcct tcagcccggg cctttcttcc 3060
ttgtgctggg atctacagag cgaacgggag aaccctctgc ttcctcagat tcaactggga 3120
agccacgaca gcccaaccct gactggcgtt actctgcctc cctgagagca ggcatgcaca 3180
gctctgtgca cctagaggag gctggcattc tacgggctgg tccaggaggg gcctgatcag 3240
cagtggccaa ccagtatcca gtgcaacacc cagaacccag aggcaggaga agtgtcccct 3300
cccagtcggt gcgggtgtca acagcaacag cgtggacctt taaatacgga ccaggcaacc 3360
ccaaacaatc cgagtcccgg tgagttgccc gacaaattca ttatcccagg atctcctggc 3420
aatcatctcc atccggcagg agcctactta cagccacaat tgacaaaagt gacttcataa 3480
cccttcggca aaaaggagga gacccagaaa aagaagaaaa agaagaaggg ttaccagacc 3540
caggagaaaa aagagaaagg gaacagcacg acgtgacaac cagtgaccac gtgaggtcct 3600
caaatgggaa acaagccact tagccagttt tttgtaataa tgggcaaatc tctcccatgt 3660
aggcaattgc cctgctcctt gtttcctatc tacattgagc cctcttagag acccgtcaga 3720
taatctgcag ataagttccc tggtgtctgt gctagaacgg catttaacac gtttttgtcg 3780
taaaaagctt tactaagtct ggttgttaac tctttctctc cactctggct gtgttttcag 3840
aacctataaa gagcagaccc agagtgtgtc ctgttgctcc tccggccgca ataggagagg 3900
cttcccagcc ccgccagtga gaggtgtgga ctctctgccc tgtgctccgg ggatcctgtc 3960
ttcgatgaca cttgcatggg caggctgaaa agttttgaga ttgagcagct tgggagtttg 4020
tggcccaccc tgggggggtta anttgttgct tttgggctaa ccccggcggg ggtaattgcc 4080
gagtgccaga tattggctga gaccgagcca gcttagacta attgggtaca agggaaaggc 4140
aagataacac aacgacaaat aaaacagcgg aagttatcag tatggagggg aaagtgtaaa 4200
ctctaaaggc gaccagacct ttcatagaat ccttacaact caagaggtgg cagccacctc 4260
tcttaggaga caaaacgtac tcgcccacca acaagactat taggagacca ctaaaatctg 4320
ttggctagtg acgtcattat acctaaaatc tggcattcat tacctggcaa ggccaaacag 4380
ttcaggtgtt taaacagaga atacaccgct gggaaacaga agcagatctg atgtgattcg 4440
ctatacatgt gcatgtgctc actttattaa aaattctttt gcacacaatt gtttatggaa 4500
aagggccaga tccttttttcc aatacttatg gcaaaagcaa aagaaaaccc cggacacctt 4560
caccttccg ctgtttgttg tttcactagg atttatttaa aaaaagagaa agtctatagc 4620
tataaagtct ttaaagagaa atatgaatac aattccccta aactctgcct caaaagagaa 4680
ttcaggtcta caacgggcgc agtttaaaat ttggactcac ttggnctgct acacgaagtg 4740
ctcttataga gaattgcctg aaacatctgt attatatcgg ccaccctgcc caatcacagc 4800
tttactcttt caggtcatct ctggggctgc cctcttgaca tgtattacta aataaaatga 4860
tctctctttc tctcgtctcg tctctctttt ctaagaaacc aattatgtgc acctttgata 4920
ccaccaaccc ttctctaacc caacctatat atccagaccc caaaaattga agaaaaatat 4980
tggttgttct catacaggtg gagcagattt ctgcaatcta cttaattctg gtggacttgg 5040
tctgggtggt gctagccata caccttcgtc gtttggttta gttttccctt tctaaaacca 5100
cctcctgaat tgtctaattc ttaactaacc accctatgaa tgttaccccg agaatcccat 5160
ctcccacata tgtatggctg ttatggctat gcttagactc cctggaataa taacttactt 5220
ctcgtgcttg tgtaatagtg aaaggtaata gccactatta cctcagagtg aactttaagc 5280
tttattgttg aaagtgaata tcccttataa tattcccttt gtgacaacct cgtggaaaaa 5340
atggagtgag tggttttttt aaccccttggt aatacagact tttgtgtatg aaagacccca 5400
gtaaaatttc ttttttaaat ccagatactg gtgattcaag gaattttatt tatggtccag 5460
cccagagcca tctcgtgccc agacttctgc tggcaagggg agtggataaa gctgttttgg 5520
ttcttagtaa caattttgga atgaatactg acaatattcc atgaagggtg tgcaagcaca 5580
aattttacca atctgacctc ttgtgaagtt gcagtaatgc tttgaaattt ctaatgggta 5640
tcctgaaata tcagctcata ggaaagtacc aaaatttgct gtcaccttaa ataagacatt 5700
ttaattttgg ttataatgta caatttagaa agtttgatta attatattat ctatttaggc 5760
attaatataa aagaggtagg agtctgttat ttaaaaaaag ccatttaatt taaaaaaaaa 5820
ctgtctgtgt ctacttttag cttcattctc ccatattttg gaagggtgtg taaactttca 5880
agctctgcag gattgccatg gggtaaaact tgttacccaa cacatgtgaa ccatttgcta 5940
cattgtaggt tgtgatcatt ttggccccac tgaagcccca tgtatcctga cccttaacgt 6000
gccctttga actaggagaa tcgggctaat ttattaatga tgataattat aatgtatctg 6060
tacagcactt tttacatttg cgaagtgcct ttccaatcca tgttagttac tagttattac 6120

cagctgtaaa ggagttaaac acctcaagtg gaatcatttt gaaattggtg ctaattggta 6180
tttcctcctg ttatctgcta ataaatgaaa aatggtggta tgaaaaaaaa 6230

<210> 43
<211> 2940
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:007302.1:2000MAY01

<400> 43
aagaatttgg actcatatca agatgctctg aagaagaaca acccttagg atagccactg 60
caacatcatg accaaagaca aagaacctat tgttaaaagc ttccattttg tttgccttat 120
gatcataata gttggaacca gaatccagtt ctccgacgga aatgaatttg cagtagacaa 180
gtcaaaaaga ggtcttattc atgttccaaa agacctaccg ctgaaaacca aagtcttaga 240
tatgtctcag aactacatcg ctgagcttca ggtctctgac atgagctttc tatcagagtt 300
gacagttttg agactttccc ataacagaat ccagctactt gatttaagtg ttttcaagtt 360
caaccaggat ttagaatatt tggatttatc tcataatcag ttgcaaaaga tatcctgcca 420
tcctattgtg agtttcaggc atttagatct ctcattcaat gatttcaagg ccctgcccat 480
ctgtaaggaa tttggcaact tatcacaact gaatttcttg ggattgagtg ctatgaagct 540
gcaaaaatta gatttgctgc caattgctca cttgcatcta agttatatcc ttctggattt 600
aagaaattat tatataaaag aaaatgagac agaaagtcta caaattctga atgcaaaaac 660
ccttcacctt gtttttcacc caactagttt attcgctatc caagtgaaca tatcagttaa 720
tactttaggg tgcttacaac tgactaatat taaattgaat gatgacaact gtcaagtttt 780
cattaaattt ttatcagaac tcaccagagg tccaacctta ctgaatttta ccctcaacca 840
catagaaacg acttggaaat gcctggtcag agtctttcaa tttctttggc ccaaacctgt 900
ggaatatctc aatatttaca atttaacaat aattgaaagc attcgtgaag aagattttac 960
ttattctaaa acgacattga aagcattgac aatagaacat atcacgaacc aagtttttct 1020
gttttcacag acagctttgt acaccgtgtt ttctgagatg aacattatga tgttaaccat 1080
ttcagataca ccttttatac acatgctgtg tcctcatgca ccaagcacat tcaagttttt 1140
gaactttacc cagaacgttt tcacagatag tatttttgaa aaatgttcca cgttagttaa 1200
attggagaca cttatcttac aaaagaatgg attaaaagac cttttcaaag taggtctcat 1260
gacgaaggat atgccttctt tggaaatact ggatgttagc tggaattctt tggaatctgg 1320
tagacataaa gaaaactgca cttgggttga gagtatagtg gtgttaaatt tgtcttcaaa 1380
tatgcttact gactctgttt tcagatgttt acctcccagg atcaaggtac ttgatcttca 1440
cagcaataaa ataaagagcg ttcctaaaca agtcgtaaaa ctggaagctt tgcaagaact 1500
caatgttgct ttcaattctt taactgacct tcctggatgt ggcagcttta gcagcctttc 1560
tgtattgatc attgatcaca attcagtttc ccacccatcg gctgatttct tccagagctg 1620
ccagaagatg aggtcaataa aagcagggga caatccattc caatgtacct gtgagctaag 1680
agaatttgtc aaaaatatag accaagtatc aagtgaagtg ttagagggct ggcctgattc 1740
ttataagtgt gactacccag aaagttatag aggaagccca ctaaaggact ttcacatgtc 1800
tgaattatcc tgcaacataa ctctgctgat cgtcaccatc ggtgccacca tgctggtgtt 1860
ggctgtgact gtgacctccc tctgcatcta cttggatctg ccctggtatc tcaggatggt 1920
gtgccagtgg acccagactc ggcgcagggc caggaacata cccttagaag aactccaaag 1980
aaacctccag tttcatgctt ttatttcata tagtgaacat gattctgcct gggtgaaaag 2040
tgaattggta ccttacctag aaaaagaaga tatacagatt tgtcttcatg agaggaactt 2100
tgtccctggc aagagcattg tggaaaatat catcaactgc attgagaaga gttacaagtc 2160
catctttgtt ttgtctccca actttgtcca gagtgagtgg tgccattacg aactctattt 2220
tgcccatcac aatctctttc atgaaggatc taataactta atcctcatct tactggaacc 2280
cattccacag aacagcattc ccaacaagta ccacaagctg aaggctctca tgacgcagcg 2340
gacttatttg cagtggccca aggagaaaag caaacgtggg gctcttttgg gctaacatta 2400
gagccgcttt taatatgaaa ttaacactag tcactgaaaa caatgatgtg aaatcttaaa 2460
aaaatttagg aaattcaact taagaaacca ttatttactt ggatgatggt gaatagtaca 2520
gtcgtaagta actgtctgga ggtgcctcca ttatcctcat gccttcagga aagacttaac 2580
aaaaacaatg tttcatctgg ggaactgagc taggcggtga ggttagcctg ccagttagag 2640
acagcccagt ctcttctggt ttaatcatta tgtttcaaat tggaaacagt ctcttttgag 2700
taaatgctca gtttttcagc tcctctccac tctgctttcc caaatggatt ctgttgtgag 2760
caagagttta tatggcttca tggcagcaag ggaacagtca acttcagcat catatgcacc 2820
agtcctcgga gtgccctgtg aatcatattg gtctttgggt cagtgtcatc attctcttca 2880
agtctggggc ttggggaaaa aattagatca gctacggcat ataaaaaagt cttttgtttc 2940

<210> 44
<211> 4438
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:236386.4:2000MAY01

<220>
<221> unsure
<222> 3774
<223> a, t, c, g, or other

<400> 44
taagcctcag tccttgtttt cccggcctgg ctcgttgtga agccggacac atccaccctt 60
ggactcgatt caggcggctg ctgcttttct ccttgcccct cttggatttt ccggattttt 120
gaaaacccag tggcccagga gcaagaggag gaaggaggaa ggggcagatc tgcagaggaa 180
tgtgagagcc tcccaaagcg agagccgcca aaagaatctg ggagccagag ggacatccga 240
gccctgcccg ggtttctgga atggtggttt cagagtgagt ctcttctatt ttagaacgtt 300
gttccagtgg aaagtgtcga atttttcccc tcgcagggca gatttctcca ggtcacttga 360
cttttcttct gggagtagga gttaggagag attcccctct aacccccccag aggctgctaa 420
gggaggagga gactgtggac atgagccctc cctgctcaca agcatatgcc cggagacctg 480
atagggcagt ttctgggcca tggacattgc tttgaagagg gggagactgg acagcatctg 540
tgggtgctga gaccccacct taggacctga gagattgaac tgtgtaagcg ccattcagct 600
gcgagtgcat tcttggactg ccttgtgagc atccccggtc tgggcaggac cctctccttc 660
ccatctttct ataccaccca gcccagccat ggcactgaaa ggccgagccc tctatgactt 720
tcacagtgag aacaaggagg aaatcagcat ccagcaggat gaggacctgg tcatctttaa 780
cgagaactca cttggattgg ttggcttgca gggccaaaac agccgtgggg agacagggct 840
ctttcctgcc tcttatgtgg agatcgtccg ttctggcatc agcaccaacc atgctgacta 900
ctccagcagc cctgcaggct ctcccggagc ccaggtgagc ttgtacaaca gccccagtgt 960
ggccagccca gctaggagtg gtgggggcag tggcttcctc tcaaaaccag ggtagctttg 1020
aggaggatga tgatgatgac tgggatgact ggtgacgacg gatgcacagt ggtggaggag 1080
ccacagggct ggtggggctg gcgcacacaa cggggcaacc cgtcccctca accgtgtcct 1140
agcatggggc cctaccccca gcccagcaca atgcccttcc ggcccaagcc aacaatgtga 1200
ggcggcagga cagcctggca tctgccaagg cgaggcagtg tggtgggcca gtaacactca 1260
accgtttctc atgctttgtg cgttctggaa tggaagccct taatcctggg tgatgtgccc 1320
atgatggcac aagatcgctg agacatactc cattgaaatg ggccctcgtg gcccccagtg 1380
tgaaggcgca atccccaccc atttgcctgc tctgtggagg accccacaaa acagaccaaa 1440
ttcaagggca tcaaaagcta catctcctac aagctcacac ccacccatgc tgcctcaccc 1500
gtctaccggc gctacaaaca ctttgactgg ctctataacc cgcctgctac acaagttcac 1560
tgtcatctcg gtgccccacc tgcctgagaa gcaggccact ggccgcttcg aggaggactt 1620
catcgaaaag cggaagcgga gactcatcct ctggatggac cacatgacca gccaccctgt 1680
gctctcccag tacgaaggct tccagcattt cctcagctgc ctggatgaca agcagtggaa 1740
gatgggcaaa cgccgggcgg agaaggatga gatggtgggt gccagcttcc tgctcacctt 1800
ccagatcccc accgagcacc aggacttgca ggacgtggaa gatcgcgtgg acactttcaa 1860
ggccttcagt aagaagatgg acgacagcgt cctgcagctc agcactgtgg catcagagct 1920
ggtgcgtaaa catgtggggg gcttcccgca aggaattcca gaacgctggg cagtgccttc 1980
caggccatca gtcattcctt ccagatggac ccccctttt gctctgaggc cctcaacagt 2040
gccatttctc acacgggccg tacctatgaa gccatcgggg agatgtttgc tgagcagccc 2100
aagaatgacc tcttccagat gctggacaca ctgtctctct accagggcct gctctccaac 2160
ttccctgaca tcatccatct acaaaaaggc gccttcgcca aggtgaagga gagccaacgc 2220
atgagtgacg agggccgcat ggtgcaggac gaggcagacg gcattcgcag gcgctgccgc 2280
gtggtgggtt tcgccctgca ggccgagatg aaccacttcc accagcgccg tgagctcgac 2340
ttcaagcatc atgatgcaga actacttgcg ccagcagatc ctcttctacc agcgggtggg 2400
ccagcagctg cataagaccc tgcgcatgta tcacaccctc tgaccgcgtg tgcctgggct 2460
ccctccttca cctgggcctg gtcactgcag tgtactccac tttcacgacc accctatgcc 2520
agcagtgact gatgaattgg tcagcggtgg cggagataac cggcctgtcc tgcctcctgg 2580
tagaaggagc tttcaaggag tcatgggtgc ccctgggaaa ttccccactc cttagaagtg 2640
gggcacagca ggggtgagaa tagagtcagg agccctcgag gccaaggcct gggctgccgg 2700
tcagtccagt gaaggtcagg ccagggtctc agcctccct agagcctatt ttgcttgctc 2760
acctggcgca ctgtgtgcct tatccattca gcagacaccg aggcctgctg caccttggg 2820
tcggatgctg ggcaccccag ggctgtgaca tgcctgcctc ttcaggagtc ctcaagtgaa 2880
ggtcggggtc agacacagac agagtcaact gcagtactga ctgactgctt taaatgacgg 2940
gatttttgga agctctatag aagggaccac agctattcca ctggtcaggg tagactccat 3000
agagtaggct acatttgggg cagtgttttg aagaatctag caaggaccta ggcccagaca 3060
gtacatgcgg gacgaagaga cttctaccgg gagaggaaca gcatgaggcc aaagttatgg 3120
agggcttgca aacttctccc tcttctctcc ccttactttc caaggcaagt taggtgacgc 3180
tttccatggg gattctcggc ctgtgtggta aggaacgagg atctcccttg ctccccatgt 3240
agctggtctg tccgtgacat caccctgtcc cctgcaggag ggggctacag gccatctccc 3300
ttcctgtagg cctctgactc ccctcccact tttggggccc tcagcttatc tcgcgcatgg 3360
ggaccattcg cagcatcctc gccctcctgc ggactcaaga tccatgagat ataagccctg 3420

ggccccagat ccctggtgac accttccttg gagaagactc tcaaaagtga ctgtatattt 3480
gagttcacca gcaataactc cccacacttc gaagcatggt ccaaacccat ggatcccagg 3540
gtccttgggc ctctgtgggc actgtcttcc caagatcctt cctgttgcaa caatgggaaa 3600
ccttaagagg aaaaagacag gggcctgctt tgcccagccc atgcgaaggg attccatgcc 3660
cacctgccct ctgcctgcct cgctggaatg tgggcccctg ctcccccgtc agggtggtgc 3720
tgtctctgac ctatgtttac gatccccgag gggtttttgg cttccccttc ccanccaggt 3780
cagggtgtgg ttccagcagc ttgctgtggg gtgctgacat gtgtcaccac tgcccccctt 3840
gtcccccggg ggggtcatgg tctcctcctg gatgctgctc cttgaatctt ttttcttgat 3900
aaaccttta caattaagat aacacaagca tgacttttc tgtttggatc ccagaaaggc 3960
ggagggcagg agaaggatag agccctaatt gctcctgaga gccattggat gagattctga 4020
ggtcgtggtg ggcacaaatt ttccacagaa cctcaaaagt tcaggggagg gctatgctgg 4080
tggaaggtgc cagcaggcag gaggagctag aggcggctgt ggacccctgg gtggatccat 4140
ccctccctag aacgcactct tgtctctaaa acaggtggag tgctgcccag gggactggct 4200
gtactgcctt gtgatctggg gctgagggtt gtatgaggaa gggacaggac gctgtgccct 4260
aggacaatta atagatggtg gctcctctcc ccaaggagcc atgccctggc cttgcccttg 4320
aaaagcccta gtccagggga gggaagtggg ggactcagaa gctgtgtctc ttccccaaac 4380
cgtcctgggt acccagccct gcggaggtcc cacattggaa ctgaagagga cgctggct 4438

<210> 45
<211> 987
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:252904.5:2000MAY01

<400> 45
cccgacttca gccccagcca gatcccgcgt caacggaggc ggaacggcgg accccgtacc 60
ctggcagcat cggagcaccg gcgggtgaag gcaaggtccc tggactggtc atatacctct 120
tgtggccctg gcagaatcaa gatgaggccc tgtcatgcct ccccagtgag gcctacagtc 180
tgagcagaca gcatggcctg ccactggcag tgaacaccat gtctgcagga ggtggccggg 240
cctttgcttg atggtatggt gtatgctctg gggggaatgg gccctgacac ggccccccag 300
gcccaggtac gtgtgtatga gccccgtcgg gactgctggc tttcgctacc ctccatgccc 360
acaccctgct atggggcctc caccttcctg cacgggaaca agatctatgt cctggggggc 420
cgccagggca agctcccggt gactgctttt gaagcctttg atctggaggc ccgtacatgg 480
acccggcatc caagcctacc cagccgtcgg gcctttgctg gctgcgccat ggctgaaggc 540
agcgtcttta gcctgggtgg cctgcagcag cctgggcccc acaacttcta ctctcgccca 600
cactttgtca acactgtgga gatgtttgac ctggagcatg ggtcctggac caaattgccc 660
cgcagcctgc gcatgaggga taagagggca gactttgtgg ttgggtccct tgggggccac 720
attgtggcca ttgggggcct tggaaaccag ccatgtcctt tgggctctgt ggagagcttt 780
agccttgcac ggcggcgctg ggaggcattg cctgccatgc ccactgcccg ctgctcctgc 840
tctagtctgc aggctgggcc ccggctgttt gttattgggg gtgtggccca gggccccagt 900
caagccgtgg aggcactgtg tctgcgtgat ggggtctgaa ggcttggtgg agctgtccac 960
tgagcagctc attggggatc cactagt 987

<210> 46
<211> 263
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:977683.1.orf3:2000FEB18

<400> 46
Gly Ser Asp Met Ala Ala Asp Leu Asn Leu Glu Trp Ile Ser Leu
1 5 10 15
Pro Arg Ser Trp Thr Tyr Gly Ile Thr Arg Gly Gly Arg Val Phe
20 25 30
Phe Ile Asn Glu Glu Ala Lys Ser Thr Thr Trp Leu His Pro Val
35 40 45
Thr Gly Glu Ala Val Val Thr Gly His Arg Arg Gln Ser Thr Asp
50 55 60
Leu Pro Thr Gly Trp Glu Glu Ala Tyr Thr Phe Glu Gly Ala Arg
65 70 75
Tyr Tyr Ile Asn His Asn Glu Arg Lys Val Thr Cys Lys His Pro

```
                 80                  85                  90
Val Thr Gly Gln Pro Ser Gln Asp Asn Cys Ile Phe Val Val Asn
                 95                 100                 105
Glu Gln Thr Val Ala Thr Met Thr Ser Glu Glu Lys Lys Glu Arg
                110                 115                 120
Pro Ile Ser Met Ile Asn Glu Ala Ser Asn Tyr Asn Val Thr Ser
                125                 130                 135
Asp Tyr Ala Val His Pro Met Ser Pro Val Gly Arg Thr Ser Arg
                140                 145                 150
Ala Ser Lys Lys Val His Asn Phe Gly Lys Arg Ser Asn Ser Ile
                155                 160                 165
Lys Arg Asn Pro Asn Ala Pro Val Val Arg Arg Gly Trp Leu Tyr
                170                 175                 180
Lys Gln Asp Ser Thr Gly Met Lys Leu Trp Lys Lys Arg Trp Phe
                185                 190                 195
Val Leu Ser Asp Leu Cys Leu Phe Tyr Tyr Arg Asp Glu Lys Glu
                200                 205                 210
Glu Gly Ile Leu Gly Ser Ile Leu Leu Pro Ser Phe Gln Ile Ser
                215                 220                 225
Phe Ala Tyr Pro Leu Lys Ile Thr Leu Ile Ala Asn Met Leu Leu
                230                 235                 240
Arg Gln Pro Ile Gln Thr Cys Gly Pro Ile Ile Ser Ala Leu Ile
                245                 250                 255
Gln Glu Arg Lys Trp Ser Cys Gly
                260
```

<210> 47
<211> 217
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:893050.1.orf1:2000FEB18

<400> 47

```
Ser Leu Pro Ser Thr Ser Phe Arg Val Ser Ser Leu Phe Ser Gly
  1                 5                  10                  15
His Leu Glu Val Leu Lys Leu Leu Val Ala Arg Gly Ala Asp Leu
                 20                  25                  30
Gly Cys Lys Ala Arg Lys Gly Tyr Gly Leu Leu His Thr Ala Ala
                 35                  40                  45
Ala Ser Gly Gln Ile Glu Val Val Lys Tyr Leu Leu Arg Met Gly
                 50                  55                  60
Ala Glu Ile Asp Glu Pro Asn Ala Phe Gly Asn Thr Ala Leu His
                 65                  70                  75
Ile Ala Cys Tyr Leu Gly Gln Asp Ala Val Ala Ile Glu Leu Val
                 80                  85                  90
Asn Ala Gly Ala Asn Val Asn Gln Pro Asn Asp Lys Gly Phe Thr
                 95                 100                 105
Pro Leu His Val Ala Ala Val Ser Thr Asn Gly Ala Leu Cys Leu
                110                 115                 120
Glu Leu Leu Val Asn Asn Gly Ala Asp Val Asn Tyr Gln Ser Lys
                125                 130                 135
Glu Gly Lys Ser Pro Leu His Met Ala Ala Ile His Gly Arg Phe
                140                 145                 150
Thr Arg Ser Gln Ile Leu Ile Gln Asn Gly Ser Glu Ile Asp Cys
                155                 160                 165
Ala Asp Lys Phe Gly Asn Thr Pro Leu His Val Ala Ala Arg Tyr
                170                 175                 180
Gly His Glu Leu Leu Ile Ser Thr Leu Met Thr Asn Gly Ala Asp
                185                 190                 195
Thr Gly Arg Arg Gly Ile His Asp Met Phe Pro Leu His Leu Ala
                200                 205                 210
Val Leu Phe Gly Phe Ser Asp
                215
```

<210> 48
<211> 716
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:980153.1.orf1:2000FEB18

<220>
<221> unsure
<222> 683
<223> unknown or other

<400> 48

```
Gln Arg Gly Ala Lys Thr Arg Leu Arg Pro Phe Ser Pro Arg His
  1               5                  10                  15
Cys Tyr Lys Ala Ala Thr Ile Lys Asp Val Phe Gly Arg Asn Ala
            20                  25                  30
Leu His Pro Cys Phe Leu Leu Val Glu Lys Lys Gly Val Leu Asp
            35                  40                  45
Trp Leu Ile Gln Lys Gly Val Asp Leu Leu Val Lys Asp Lys Glu
            50                  55                  60
Ser Gly Trp Thr Ala Leu His Arg Ser Ile Phe Tyr Gly His Ile
            65                  70                  75
Asp Cys Val Trp Ser Leu Leu Lys His Gly Val Ser Leu Tyr Ile
            80                  85                  90
Gln Asp Lys Glu Gly Leu Ser Ala Leu Asp Leu Val Met Lys Asp
            95                 100                 105
Arg Pro Thr His Val Val Phe Lys Asn Thr Asp Pro Thr Asp Val
           110                 115                 120
Tyr Thr Trp Gly Asp Asn Thr Asn Phe Thr Leu Gly His Gly Ser
           125                 130                 135
Gln Asn Ser Lys His His Pro Glu Leu Val Asp Leu Phe Ser Arg
           140                 145                 150
Ser Gly Ile Tyr Ile Lys Gln Val Val Leu Cys Lys Phe His Ser
           155                 160                 165
Val Phe Leu Ser Gln Lys Gly Gln Val Tyr Thr Cys Gly His Gly
           170                 175                 180
Pro Gly Gly Arg Leu Gly His Gly Asp Glu Gln Thr Cys Leu Val
           185                 190                 195
Pro Arg Leu Val Glu Gly Leu Asn Gly His Asn Cys Ser Gln Val
           200                 205                 210
Ala Ala Ala Lys Asp His Thr Val Val Leu Thr Glu Asp Gly Cys
           215                 220                 225
Val Tyr Thr Phe Gly Leu Asn Ile Phe His Gln Leu Gly Ile Ile
           230                 235                 240
Pro Pro Pro Ser Ser Cys Asn Val Pro Arg Gln Ile Gln Ala Lys
           245                 250                 255
Tyr Leu Lys Gly Arg Thr Ile Ile Gly Val Ala Ala Gly Arg Phe
           260                 265                 270
His Thr Val Leu Trp Thr Arg Glu Ala Val Tyr Thr Met Gly Leu
           275                 280                 285
Asn Gly Gly Gln Leu Gly Cys Leu Leu Asp Pro Asn Gly Glu Lys
           290                 295                 300
Cys Val Thr Ala Pro Arg Gln Val Ser Ala Leu His His Lys Asp
           305                 310                 315
Ile Ala Leu Ser Leu Val Ala Ala Ser Asp Gly Ala Thr Val Cys
           320                 325                 330
Val Thr Thr Arg Gly Asp Ile Tyr Leu Leu Ala Asp Tyr Gln Cys
           335                 340                 345
Lys Lys Met Ala Ser Lys Gln Leu Asn Leu Lys Lys Val Leu Val
           350                 355                 360
Ser Gly Gly His Met Glu Tyr Lys Val Asp Pro Glu His Leu Lys
           365                 370                 375
Glu Asn Gly Gly Gln Lys Ile Cys Ile Leu Ala Met Asp Gly Ala
           380                 385                 390
```

```
Gly Arg Val Phe Cys Trp Arg Ser Val Asn Ser Ser Leu Lys Gln
                395                 400                 405
Cys Arg Trp Ala Tyr Pro Arg Gln Val Phe Ile Ser Asp Ile Ala
                410                 415                 420
Leu Asn Arg Asn Glu Ile Leu Phe Val Thr Gln Asp Gly Glu Gly
                425                 430                 435
Phe Arg Gly Arg Trp Phe Glu Glu Lys Arg Lys Ser Ser Glu Lys
                440                 445                 450
Lys Glu Ile Leu Ser Asn Leu His Asn Ser Ser Ser Asp Val Ser
                455                 460                 465
Tyr Val Ser Asp Ile Asn Ser Val Tyr Glu Arg Ile Arg Leu Glu
                470                 475                 480
Lys Leu Thr Phe Ala His Arg Ala Val Ser Val Ser Thr Asp Pro
                485                 490                 495
Ser Gly Cys Asn Phe Ala Ile Leu Gln Ser Asp Pro Lys Thr Ser
                500                 505                 510
Leu Tyr Glu Ile Pro Ala Val Ser Ser Ser Ser Phe Phe Glu Glu
                515                 520                 525
Phe Gly Lys Leu Leu Arg Glu Ala Asp Glu Met Asp Ser Ile His
                530                 535                 540
Asp Val Thr Phe Gln Val Gly Asn Arg Leu Phe Pro Ala His Lys
                545                 550                 555
Tyr Ile Leu Ala Val His Ser Asp Phe Phe Gln Lys Leu Phe Leu
                560                 565                 570
Ser Asp Gly Asn Thr Ser Glu Phe Thr Asp Ile Tyr Gln Lys Asp
                575                 580                 585
Glu Asp Ser Ala Gly Cys His Leu Phe Val Val Glu Lys Val His
                590                 595                 600
Pro Asp Met Phe Glu Tyr Leu Leu Gln Phe Ile Tyr Thr Asp Thr
                605                 610                 615
Cys Asp Phe Leu Thr His Gly Phe Lys Pro Arg Ile His Leu Asn
                620                 625                 630
Lys Asn Pro Glu Glu Tyr Gln Gly Thr Leu Asn Ser His Leu Asn
                635                 640                 645
Lys Val Asn Phe His Glu Asp Asp Asn Gln Lys Ser Ala Phe Glu
                650                 655                 660
Val Tyr Lys Ser Asn Gln Ala Gln Thr Val Ser Glu Arg Gln Lys
                665                 670                 675
Ser Lys Pro Lys Ser Cys Lys Xaa Gly Lys Asn Ile Arg Glu Asp
                680                 685                 690
Asp Pro Val Arg Met Leu Gln Thr Val Ala Lys Lys Phe Asp Phe
                695                 700                 705
Ser Asn Leu Ser Ser Arg Leu Asp Gly Val Arg
                710                 715

<210> 49
<211> 107
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:350398.1.orf3:2000FEB18

<220>
<221> unsure
<222> 22
<223> unknown or other

<400> 49
Glu Pro Leu Ser Pro Pro Gly Arg Ile Pro Gly Ala Ala Gly Glu
  1               5                  10                  15
Cys Glu Gly Pro Gln Gly Xaa Phe Ala Ser Arg Gln Pro Tyr Ser
                20                  25                  30
Arg Phe Leu Leu Arg Tyr Trp His Leu Thr Pro Ile Thr Pro Trp
                35                  40                  45
Ala Ile Val Pro Val Trp Ser Pro Arg Gly Arg Ser Arg Gly Ser
```

```
                50                  55                  60
Pro Asn Ser Thr Ser Gln Thr Ser Ile Gln Ala Gly Thr Ser Thr
                65                  70                  75
Leu Leu Ala Ser Arg His Gln Asn Ile Trp Glu Asp Met Cys Val
                80                  85                  90
Ser Thr Cys Met Trp Gly His Thr Gly Gly Asn Met Gly Met Arg
                95                 100                 105
Ala Val
```

<210> 50
<211> 645
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:475551.1.orf3:2000FEB18

<220>
<221> unsure
<222> 141
<223> unknown or other

<400> 50

```
Leu Gln Gly Gln Ser Gly Ala Asp Met Asp Lys Arg Val Lys Lys
  1               5                  10                  15
Leu Pro Leu Met Ala Leu Ser Thr Thr Met Ala Glu Ser Phe Lys
                20                  25                  30
Glu Leu Asp Pro Asp Ser Ser Met Gly Lys Ala Leu Glu Met Ser
                35                  40                  45
Cys Ala Ile Gln Asn Gln Leu Ala Arg Ile Leu Ala Glu Phe Glu
                50                  55                  60
Met Thr Leu Glu Arg Asp Val Leu Gln Pro Leu Ser Arg Leu Ser
                65                  70                  75
Glu Glu Glu Leu Pro Ala Ile Leu Lys His Lys Lys Ser Leu Gln
                80                  85                  90
Lys Leu Val Ser Asp Trp Asn Thr Leu Lys Asn Arg Leu Ser Gln
                95                 100                 105
Ala Thr Lys Asn Ser Gly Ser Ser Gln Gly Leu Gly Gly Ser Pro
               110                 115                 120
Gly Ser His Ser His Thr Thr Met Ala Asn Lys Val Glu Thr Leu
               125                 130                 135
Phe Tyr Cys Ser Arg Xaa Ser Pro Arg Lys Val Glu Gln Cys Arg
               140                 145                 150
Asp Glu Tyr Leu Ala Asp Leu Tyr His Phe Val Thr Lys Glu Asp
               155                 160                 165
Ser Tyr Ala Asn Tyr Phe Ile Arg Leu Leu Glu Ile Gln Ala Asp
               170                 175                 180
Tyr His Arg Arg Ser Leu Ser Ser Leu Asp Thr Ala Leu Ala Glu
               185                 190                 195
Leu Arg Glu Asn His Gly Gln Ala Asp His Ser Pro Ser Met Thr
               200                 205                 210
Ala Thr His Phe Pro Arg Val Tyr Gly Val Ser Leu Ala Thr His
               215                 220                 225
Leu Gln Glu Leu Gly Arg Glu Ile Ala Leu Pro Ile Glu Ala Cys
               230                 235                 240
Val Met Met Leu Leu Ser Glu Gly Met Lys Glu Glu Gly Leu Phe
               245                 250                 255
Arg Leu Ala Ala Gly Ala Ser Val Leu Lys Arg Leu Lys Gln Thr
               260                 265                 270
Met Ala Ser Asp Pro His Ser Leu Glu Glu Phe Cys Ser Asp Pro
               275                 280                 285
His Ala Val Ala Gly Ala Leu Lys Ser Tyr Leu Arg Glu Leu Pro
               290                 295                 300
Glu Pro Leu Met Thr Phe Asp Leu Tyr Asp Asp Trp Met Arg Ala
               305                 310                 315
```

```
Ala Ser Leu Lys Glu Pro Gly Ala Arg Leu Gln Ala Leu Gln Glu
                320                 325                 330
Val Cys Ser Arg Leu Pro Pro Glu Asn Leu Ser Asn Leu Arg Tyr
                335                 340                 345
Leu Met Lys Phe Leu Ala Arg Leu Ala Glu Glu Gln Glu Val Asn
                350                 355                 360
Lys Met Thr Pro Ser Asn Ile Ala Ile Val Leu Gly Pro Asn Leu
                365                 370                 375
Leu Trp Pro Pro Glu Lys Glu Gly Asp Gln Ala Gln Leu Asp Ala
                380                 385                 390
Ala Ser Val Ser Ser Ile Gln Val Val Gly Val Val Glu Ala Leu
                395                 400                 405
Ile Gln Ser Ala Asp Thr Leu Phe Pro Gly Asp Ile Asn Phe Asn
                410                 415                 420
Val Ser Gly Leu Phe Ser Ala Val Thr Leu Gln Asp Thr Val Ser
                425                 430                 435
Asp Arg Leu Ala Ser Glu Glu Leu Pro Ser Thr Ala Val Pro Thr
                440                 445                 450
Pro Ala Thr Thr Pro Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro
                455                 460                 465
Ala Pro Ala Leu Ala Ser Ala Ala Thr Lys Glu Arg Thr Glu Ser
                470                 475                 480
Glu Val Pro Pro Arg Pro Ala Ser Pro Lys Val Thr Arg Ser Pro
                485                 490                 495
Pro Glu Thr Ala Ala Pro Val Glu Asp Met Ala Arg Arg Thr Lys
                500                 505                 510
Arg Pro Ala Pro Ala Arg Pro Thr Met Pro Pro Pro Gln Val Ser
                515                 520                 525
Gly Ser Arg Ser Ser Pro Pro Ala Pro Pro Leu Pro Pro Gly Ser
                530                 535                 540
Gly Ser Pro Gly Thr Pro Gln Ala Leu Pro Arg Arg Leu Val Gly
                545                 550                 555
Ser Ser Leu Arg Ala Pro Thr Val Pro Pro Pro Leu Pro Pro Thr
                560                 565                 570
Pro Pro Gln Pro Ala Arg Arg Gln Ser Arg Arg Ser Pro Ala Ser
                575                 580                 585
Pro Ser Pro Ala Ser Pro Gly Pro Ala Ser Pro Ser Pro Val Ser
                590                 595                 600
Leu Ser Asn Pro Ala Gln Val Asp Leu Gly Ala Ala Thr Ala Glu
                605                 610                 615
Gly Gly Ala Pro Glu Ala Ile Ser Gly Val Pro Thr Pro Pro Ala
                620                 625                 630
Ile Pro Pro Gln Pro Arg Pro Arg Ser Leu Ala Ser Glu Thr Asn
                635                 640                 645


<210> 51
<211> 177
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:481407.2.orf3:2000FEB18

<400> 51
Cys Gln Gly Arg Cys Glu Arg Leu Arg Arg Val Gly Val Glu Pro
  1                 5                  10                  15
Gln Leu Ser Arg Gly Leu Ala Leu Phe Trp Ser Pro Arg Pro Asn
                 20                  25                  30
Pro Pro Glu Glu Met Ser Gly Gly Leu Ala Pro Ser Lys Ser Thr
                 35                  40                  45
Val Tyr Val Ser Asn Leu Pro Phe Ser Leu Thr Asn Asn Asp Leu
                 50                  55                  60
Tyr Arg Ile Phe Ser Lys Tyr Gly Lys Val Val Lys Val Thr Ile
                 65                  70                  75
Met Lys Asp Lys Asp Thr Arg Lys Ser Lys Gly Val Ala Phe Ile
```

```
                  80                  85                  90
Leu Phe Leu Asp Lys Asp Ser Ala Gln Asn Cys Thr Arg Ala Ile
                  95                 100                 105
Asn Asn Lys Gln Leu Phe Gly Arg Val Ile Lys Ala Ser Ile Ala
                 110                 115                 120
Ile Asp Asn Gly Arg Ala Ala Glu Phe Ile Arg Arg Arg Asn Tyr
                 125                 130                 135
Phe Asp Lys Ser Lys Cys Tyr Glu Cys Gly Glu Ser Gly His Leu
                 140                 145                 150
Ser Tyr Ala Cys Pro Lys Asn Met Leu Gly Glu Arg Glu Pro Pro
                 155                 160                 165
Lys Lys Lys Glu Lys Lys Glu Lys Lys Glu Ser Ser
                 170                 175

<210> 52
<211> 217
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:443580.1.orf1:2000FEB01

<400> 52
Glu Thr Ser Leu Arg Ser Gly Gln Ile Pro Thr Leu Asp Ser Ser
  1                 5                  10                  15
Glu His Asn Leu Ser Pro Glu Pro Leu Glu Leu Asp Arg Met Pro
                  20                  25                  30
His Ser Pro Leu Ile Ser Ile Pro His Val Trp Cys His Pro Glu
                  35                  40                  45
Glu Glu Glu Arg Met His Asp Glu Leu Leu Gln Ala Val Ser Lys
                  50                  55                  60
Gly Pro Val Met Phe Arg Asp Val Ser Ile Asp Phe Ser Gln Glu
                  65                  70                  75
Glu Trp Glu Cys Leu Asp Ala Asp Gln Met Asn Leu Tyr Lys Glu
                  80                  85                  90
Val Met Leu Glu Asn Phe Ser Asn Leu Val Ser Val Gly Leu Ser
                  95                 100                 105
Asn Ser Lys Pro Ala Val Ile Ser Leu Leu Glu Gln Gly Lys Glu
                 110                 115                 120
Pro Trp Met Val Asp Arg Glu Leu Thr Arg Gly Leu Cys Ser Asp
                 125                 130                 135
Leu Glu Ser Met Cys Glu Thr Lys Ile Leu Ser Leu Lys Lys Arg
                 140                 145                 150
His Phe Ser Gln Val Ile Ile Thr Arg Glu Asp Met Ser Thr Phe
                 155                 160                 165
Ile Gln Pro Thr Phe Leu Ile Pro Pro Gln Lys Thr Met Ser Glu
                 170                 175                 180
Glu Lys Pro Trp Glu Cys Lys Ile Cys Gly Lys Thr Phe Asn Gln
                 185                 190                 195
Asn Ser Gln Phe Ile Gln His Gln Arg Ile His Phe Gly Glu Lys
                 200                 205                 210
His Tyr Glu Ser Lys Glu Lys
                 215

<210> 53
<211> 151
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:803015.1.orf3:2000FEB01

<400> 53
Ala Gly Cys Gly Trp Asp Pro Val Phe Pro Ala Pro Arg Gly Thr
  1                 5                  10                  15
```

```
Trp Phe Leu Cys Pro Gly Phe Cys His Ser Val Thr Tyr Ala Met
                  20                  25                  30
Pro Cys Cys Ser His Arg Arg Cys Arg Glu Asp Pro Gly Thr Ser
                  35                  40                  45
Glu Ser Gln Glu Met Asp Pro Val Ala Phe Asp Asp Val Ala Val
                  50                  55                  60
Asn Phe Thr Gln Glu Glu Trp Ala Leu Asp Ile Ser Gln Arg
                  65                  70                  75
Lys Leu Tyr Lys Glu Val Met Leu Glu Thr Phe Arg Asn Leu Thr
                  80                  85                  90
Ser Val Gly Lys Ser Trp Lys Asp Gln Asn Ile Glu Tyr Glu Tyr
                  95                 100                 105
Gln Asn Pro Arg Arg Asn Phe Arg Ser Leu Ile Glu Lys Lys Val
                 110                 115                 120
Asn Glu Ile Lys Asp Asp Ser His Cys Gly Glu Thr Phe Thr Gln
                 125                 130                 135
Val Pro Asp Asp Arg Leu Asn Phe Gln Glu Lys Lys Ala Ser Pro
                 140                 145                 150
Glu
```

<210> 54
<211> 193
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:027410.3.orf3:2000MAY19

<400> 54

```
His Thr Glu Ala Arg Pro Pro Arg Arg Glu Ser Trp Ile Ser Asp
  1                 5                  10                  15
Ile Arg Ala Gly Thr Ala Pro Ser Cys Arg Asn His Ile Lys Ser
                  20                  25                  30
Ser Cys Ser Leu Ile Ala Phe Asn Ser Asp Arg Pro Gly Val Leu
                  35                  40                  45
Gly Ile Val Pro Leu Gln Gly Gln Gly Glu Asp Lys Arg Arg Val
                  50                  55                  60
Ala His Leu Gly Cys His Ser Asp Leu Val Thr Asp Leu Asp Phe
                  65                  70                  75
Ser Pro Phe Asp Asp Phe Leu Leu Ala Thr Gly Ser Ala Asp Arg
                  80                  85                  90
Thr Val Lys Leu Trp Arg Leu Pro Gly Pro Gly Gln Ala Leu Pro
                  95                 100                 105
Ser Ala Pro Gly Val Val Leu Gly Pro Glu Asp Leu Pro Val Glu
                 110                 115                 120
Val Leu Gln Phe His Pro Thr Ser Asp Gly Ile Leu Ser Trp Gln
                 125                 130                 135
Pro Met Gly Thr Trp Cys Arg Ala Pro Ser Gly Ala Glu Met Glu
                 140                 145                 150
Pro Trp Trp Ala Arg Arg Ala Arg Thr Ser Ser Cys Gly Ser Leu
                 155                 160                 165
Thr Pro Glu Gln Ser Arg Gly Pro Leu Arg Ala Arg Arg Pro Met
                 170                 175                 180
Arg Thr Ala Gly Ile Ala Gly Trp His Gly Trp Ala Pro
                 185                 190
```

<210> 55
<211> 282
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:171377.1.orf3:2000MAY19

<400> 55
Arg Pro Gln Pro Leu Arg Ala Arg Thr Ala Ala Pro Pro Arg Pro
1 5 10 15
Ser Gln Pro Ala Ser Gln Thr Gly Leu Arg Pro Thr Asp Gly Arg
20 25 30
Ser Arg Ser Gly Pro Ala Arg Leu Leu Cys Pro Gly Pro Ala Ala
35 40 45
Pro Arg Ser Pro Ala Val Ser Ala Ala Ser Arg Pro Glu Ser Gln
50 55 60
Ala Pro Thr Pro Arg Pro Ala Val Ala Ala Pro Ser Met Ser Ser
65 70 75
Thr Glu Arg Arg Pro Ala Gly Arg Arg Asp Arg Ser Pro Arg Gln
80 85 90
Gln Val Asp Arg Leu Leu Val Gly Leu Arg Trp Arg Arg Leu Glu
95 100 105
Glu Pro Leu Gly Phe Ile Lys Val Leu Gln Trp Leu Phe Ala Ile
110 115 120
Phe Ala Phe Gly Ser Cys Gly Ser Tyr Ser Gly Glu Thr Gly Ala
125 130 135
Met Val Arg Cys Asn Asn Glu Ala Lys Asp Val Ser Ser Ile Ile
140 145 150
Val Ala Phe Gly Tyr Pro Cys Arg Leu His Arg Ile Gln Tyr Glu
155 160 165
Met Pro Leu Cys Asp Glu Glu Ser Ser Ser Lys Thr Met His Leu
170 175 180
Met Gly Asp Phe Ser Ala Pro Ala Glu Phe Phe Val Thr Leu Gly
185 190 195
Ile Phe Ser Phe Phe Tyr Thr Met Ala Ala Leu Val Ile Tyr Leu
200 205 210
Arg Phe His Asn Leu Tyr Thr Glu Asn Lys Arg Phe Pro Leu Val
215 220 225
Asp Phe Cys Val Thr Val Ser Phe Thr Phe Phe Trp Leu Val Ala
230 235 240
Ala Ala Ala Trp Gly Lys Gly Leu Thr Asp Val Lys Gly Ala Thr
245 250 255
Arg Pro Ser Ser Leu Thr Ala Ala Met Ser Val Cys His Gly Glu
260 265 270
Glu Ala Val Cys Ser Ala Gly Ala Thr Pro Ser Met
275 280

<210> 56
<211> 211
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:352559.1.orf2:2000MAY19

<400> 56
Val Val Ser Ser Thr Thr Ala Ser Ala Leu Gln Ser Gln Ser Lys
1 5 10 15
Ala Leu Leu Gln Met Lys Ser Gln Glu Glu Val Glu Val Ala Gly
20 25 30
Ile Lys Leu Cys Lys Ala Met Ser Leu Gly Ser Leu Thr Phe Thr
35 40 45
Asp Val Ala Ile Asp Phe Ser Gln Asp Glu Trp Glu Trp Leu Asn
50 55 60
Leu Ala Gln Arg Ser Leu Tyr Lys Lys Val Met Leu Glu Asn Tyr
65 70 75
Arg Asn Leu Val Ser Val Gly Leu Cys Ile Ser Lys Pro Asp Val
80 85 90
Ile Ser Leu Leu Glu Gln Glu Lys Asp Pro Trp Val Ile Lys Gly
95 100 105
Gly Met Asn Arg Gly Leu Cys Pro Asp Leu Glu Cys Val Trp Val
110 115 120
Thr Lys Ser Leu Ser Leu Asn Gln Asp Ile Tyr Glu Glu Lys Leu

```
                125                 130                 135
Pro Pro Ala Ile Ile Met Glu Arg Leu Lys Ser Tyr Asp Leu Glu
                140                 145                 150
Cys Ser Thr Leu Gly Lys Asn Trp Lys Cys Glu Asp Leu Phe Glu
                155                 160                 165
Arg Glu Leu Val Asn Gln Lys Thr His Phe Arg Gln Glu Thr Ile
                170                 175                 180
Thr His Ile Asp Thr Leu Ile Glu Lys Arg Asp His Ser Asn Lys
                185                 190                 195
Ser Gly Thr Val Phe His Leu Asn Thr Leu Ser Tyr Ile Lys Gln
                200                 205                 210
Ile
```

<210> 57
<211> 366
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:247384.1.orf2:2000MAY19

<400> 57

```
Arg Arg Gln Leu Gly Val Ala Leu Ile Pro Ser His Arg Met Asp
  1                 5                  10                  15
Tyr Lys Ser Ser Leu Ile Gln Asp Gly Asn Pro Met Glu Asn Leu
                 20                  25                  30
Glu Lys Gln Leu Ile Cys Pro Ile Cys Leu Glu Met Phe Thr Lys
                 35                  40                  45
Pro Val Val Ile Leu Pro Cys Gln His Asn Leu Cys Arg Lys Cys
                 50                  55                  60
Ala Asn Asp Ile Phe Gln Ala Ser Asn Pro Tyr Leu Pro Thr Arg
                 65                  70                  75
Gly Gly Thr Thr Met Ala Ser Gly Gly Arg Phe Arg Cys Pro Ser
                 80                  85                  90
Cys Arg His Glu Val Val Leu Asp Arg His Gly Val Tyr Gly Leu
                 95                 100                 105
Gln Arg Asn Leu Leu Val Glu Asn Ile Ile Asp Ile Tyr Lys Gln
                110                 115                 120
Glu Cys Ser Ser Arg Pro Leu Gln Lys Gly Ser His Pro Met Cys
                125                 130                 135
Lys Glu His Glu Asp Glu Lys Ile Asn Ile Tyr Cys Leu Thr Cys
                140                 145                 150
Glu Val Pro Thr Cys Ser Met Cys Lys Val Phe Gly Ile His Lys
                155                 160                 165
Ala Cys Glu Val Ala Pro Leu Gln Ser Val Phe Gln Gly Gln Lys
                170                 175                 180
Thr Glu Leu Asn Asn Cys Ile Ser Met Leu Val Ala Gly Asn Asp
                185                 190                 195
Arg Val Gln Thr Ile Ile Thr Gln Leu Glu Asp Ser Arg Arg Val
                200                 205                 210
Thr Lys Glu Asn Ser His Gln Val Lys Glu Glu Leu Ser Gln Lys
                215                 220                 225
Phe Asp Thr Leu Tyr Ala Ile Leu Asp Glu Lys Lys Ser Glu Leu
                230                 235                 240
Leu Gln Arg Ile Thr Gln Glu Gln Glu Lys Lys Leu Ser Phe Ile
                245                 250                 255
Glu Ala Leu Ile Gln Gln Tyr Gln Glu Gln Leu Asp Lys Ser Thr
                260                 265                 270
Lys Leu Val Glu Thr Ala Ile Gln Ser Leu Asp Glu Pro Gly Gly
                275                 280                 285
Ala Thr Phe Leu Leu Thr Ala Lys Gln Leu Ile Lys Ser Ile Val
                290                 295                 300
Glu Ala Ser Lys Gly Cys Gln Leu Gly Lys Thr Glu Gln Gly Phe
                305                 310                 315
Glu Asn Met Asp Phe Phe Thr Leu Asp Leu Glu His Ile Ala Asp
```

```
                320                 325                 330
Ala Leu Arg Ala Ile Asp Phe Gly Thr Asp Glu Glu Glu Glu Glu
                335                 340                 345
Phe Ile Glu Glu Glu Asp Gln Glu Glu Glu Glu Ser Thr Glu Gly
                350                 355                 360
Lys Glu Glu Gly His Gln
                365
```

<210> 58
<211> 326
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:403872.1.orf3:2000MAY19

<220>
<221> unsure
<222> 294
<223> unknown or other

<400> 58

```
Glu Met Ala Val Gly Asn Asn Thr Gln Arg Ser Tyr Ser Ile Ile
  1               5                  10                  15
Pro Cys Phe Ile Phe Val Glu Leu Val Ile Met Ala Gly Thr Val
                20                  25                  30
Leu Leu Ala Tyr Tyr Phe Glu Cys Thr Asp Thr Phe Gln Val His
                35                  40                  45
Ile Gln Gly Phe Phe Cys Gln Asp Gly Asp Leu Met Lys Pro Tyr
                50                  55                  60
Pro Gly Thr Glu Glu Glu Ser Phe Ile Thr Pro Leu Val Leu Tyr
                65                  70                  75
Cys Val Leu Ala Ala Thr Pro Thr Ala Ile Ile Phe Ile Gly Glu
                80                  85                  90
Ile Ser Met Tyr Phe Ile Lys Ser Thr Arg Glu Ser Leu Ile Ala
                95                 100                 105
Gln Glu Lys Thr Ile Leu Thr Gly Glu Cys Cys Tyr Leu Asn Pro
               110                 115                 120
Leu Leu Arg Arg Ile Ile Arg Phe Thr Gly Val Phe Ala Phe Gly
               125                 130                 135
Leu Phe Ala Thr Asp Ile Phe Val Asn Ala Gly Gln Val Val Thr
               140                 145                 150
Gly His Leu Thr Pro Tyr Phe Leu Thr Val Cys Lys Pro Asn Tyr
               155                 160                 165
Thr Ser Ala Asp Cys Gln Ala His His Gln Phe Ile Asn Asn Gly
               170                 175                 180
Asn Ile Cys Thr Gly Asp Leu Glu Val Ile Glu Lys Ala Arg Arg
               185                 190                 195
Ser Phe Pro Ser Lys His Ala Ala Leu Ser Ile Tyr Ser Ala Leu
               200                 205                 210
Tyr Ala Thr Met Tyr Ile Thr Ser Thr Ile Lys Thr Lys Ser Ser
               215                 220                 225
Arg Leu Ala Lys Pro Val Leu Cys Leu Gly Thr Leu Cys Thr Ala
               230                 235                 240
Phe Leu Thr Gly Leu Asn Arg Val Ser Glu Tyr Arg Asn His Cys
               245                 250                 255
Ser Asp Val Ile Ala Gly Phe Ile Leu Gly Thr Ala Val Ala Leu
               260                 265                 270
Phe Leu Gly Met Cys Val Val His Asn Phe Lys Gly Thr Gln Gly
               275                 280                 285
Ser Pro Ser Lys Pro Lys Pro Glu Xaa Pro Arg Gly Val Pro Leu
               290                 295                 300
Met Ala Phe Pro Arg Ile Glu Ser Pro Leu Glu Thr Leu Ser Ala
               305                 310                 315
Gln Asn His Ser Ala Ser Met Thr Glu Val Thr
               320                 325
```

<210> 59
<211> 156
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1135213.1.orf1:2000MAY19

<400> 59
Leu Cys Gly Asp Tyr Ser Cys Leu Thr Thr Glu Phe Pro Thr Glu
1 5 10 15
Ile Met Glu Glu Lys Gln Gln Ile Ile Leu Ala Asn Gln Asp Gly
20 25 30
Gly Thr Val Ala Gly Ala Ala Pro Thr Phe Phe Val Ile Leu Lys
35 40 45
Gln Pro Gly Asn Gly Lys Thr Asp Gln Gly Ile Leu Val Thr Asn
50 55 60
Gln Asp Ala Cys Ala Leu Ala Ser Ser Val Ser Ser Pro Val Lys
65 70 75
Ser Lys Gly Lys Ile Cys Leu Pro Ala Asp Cys Thr Val Gly Gly
80 85 90
Ile Thr Val Thr Leu Asp Asn Asn Ser Met Trp Asn Glu Phe Tyr
95 100 105
His Arg Ser Thr Glu Met Ile Leu Thr Lys Gln Gly Arg Arg Met
110 115 120
Phe Pro Tyr Cys Arg Tyr Trp Ile Thr Gly Leu Asp Ser Asn Leu
125 130 135
Lys Tyr Ile Leu Val Met Asp Ile Ser Pro Val Asp Asn His Arg
140 145 150
Tyr Lys Trp Asn Gly Arg
155

<210> 60
<211> 262
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:474284.2.orf2:2000MAY19

<400> 60
Ser Ser Pro Thr Ser Trp Arg Ser Ser Met Pro Cys Thr Trp Arg
1 5 10 15
Ser Arg Arg Arg Arg Cys Thr Ala Cys Ser Ala Ala Ala Ala Pro
20 25 30
Pro Leu Pro Ala Gln Lys Val Cys Leu Arg Cys Glu Ala Pro Cys
35 40 45
Cys Gln Ser His Val Gln Thr His Leu Gln Gln Pro Ser Thr Ala
50 55 60
Arg Gly His Leu Leu Val Glu Ala Asp Asp Val Arg Ala Trp Ser
65 70 75
Cys Pro Gln His Asn Ala Tyr Arg Leu Tyr His Cys Glu Ala Glu
80 85 90
Gln Val Ala Val Cys Gln Tyr Cys Cys Tyr Tyr Ser Gly Ala His
95 100 105
Gln Gly His Ser Val Cys Asp Val Glu Ile Arg Arg Asn Glu Ile
110 115 120
Arg Lys Met Leu Met Lys Gln Gln Asp Arg Leu Glu Glu Arg Glu
125 130 135
Gln Asp Ile Glu Asp Gln Leu Tyr Lys Leu Glu Ser Asp Lys Arg
140 145 150
Leu Val Glu Glu Lys Val Asn Gln Leu Lys Glu Glu Val Arg Leu
155 160 165
Gln Tyr Glu Lys Leu His Gln Leu Leu Asp Glu Asp Leu Arg Gln
170 175 180

```
Thr Val Glu Val Leu Asp Lys Ala Gln Ala Lys Phe Cys Ser Glu
                185                 190                 195
Asn Ala Ala Gln Ala Leu His Leu Gly Glu Arg Met Gln Glu Ala
                200                 205                 210
Lys Lys Leu Leu Gly Ser Leu Gln Leu Leu Phe Asp Lys Thr Glu
                215                 220                 225
Asp Val Ser Phe Met Lys Asn Thr Lys Ser Val Lys Ile Leu Met
                230                 235                 240
Asp Ser Arg Cys Pro Val His Trp Pro Gln Asp Pro Asp Leu His
                245                 250                 255
Glu Gln Gln Pro Phe Pro His
                260
```

<210> 61
<211> 132
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:342147.1.orf3:2000MAY19

<400> 61

```
Lys Thr Asn Leu Tyr Cys Ser Pro Tyr Phe Ile Asp Cys Asn Arg
  1                 5                  10                  15
Ser Ile Glu Val Thr Phe Ile Leu Ser Trp Ile Val Cys Ser Tyr
                 20                  25                  30
Ala Val Cys Lys Glu Arg Asn Gly Met Gly Gly Cys Glu Lys Glu
                 35                  40                  45
Glu Leu Val Val Asp Phe Gly Gly Ala Gly Trp Arg Ser Leu Cys
                 50                  55                  60
Leu Cys Ser Arg Leu Gly Cys Ala Ala Pro Arg Pro Arg Cys Pro
                 65                  70                  75
Asp Phe Arg Arg Pro Asp Ala Ser Leu Thr Ser Ala Ser Ala Arg
                 80                  85                  90
Gly Cys Trp Arg Pro Ser Trp Leu Arg Ser Ala Pro Pro Arg Ser
                 95                 100                 105
Pro Pro Thr Thr Cys Ala His Pro Ala Trp Arg Cys Pro Ser Pro
                110                 115                 120
Arg Cys Arg Arg Thr Pro Ala Pro Phe Arg Cys Cys
                125                 130
```

<210> 62
<211> 167
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1097300.1.orf2:2000MAY19

<400> 62

```
Pro Pro Arg Arg Arg Pro Cys Trp Phe Leu Cys Gly Leu Leu Ser
  1                 5                  10                  15
Arg Met Val Lys Leu Phe Ile Gly Asn Leu Pro Arg Glu Ala Thr
                 20                  25                  30
Glu Gln Glu Ile Arg Ser Leu Phe Glu Gln Tyr Gly Lys Val Leu
                 35                  40                  45
Glu Cys Asp Ile Ile Lys Asn Tyr Gly Phe Val His Ile Glu Asp
                 50                  55                  60
Lys Thr Ala Ala Glu Asp Ala Ile Arg Asn Leu His His His Lys
                 65                  70                  75
Pro His Gly Val Asn Ile Asn Ala Glu Ala Ser Lys Asn Lys Ser
                 80                  85                  90
Lys Ala Pro Thr Lys Leu His Val Gly Asn Ile Ser Pro Thr Cys
                 95                 100                 105
Thr Asn Gln Glu Leu Arg Ala Lys Phe Glu Glu His Gly Pro Ala
```

```
                110                 115                 120
Ile Glu Cys Asp Ile Ala Lys Asp Tyr Ala Phe Ala His Met Glu
                125                 130                 135
Arg Ala Glu Asp Ala Ala Glu Ala Ile Arg Gly Leu Asp Asn Thr
                140                 145                 150
Glu Phe Gln Gly Glu Leu Leu Trp Ala Trp Val Val Ala Pro Ser
                155                 160                 165
Gly Val
```

<210> 63
<211> 570
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:444850.9.orf1:2000MAY19

<220>
<221> unsure
<222> 569-570
<223> unknown or other

<400> 63

```
Lys His Arg Gln Glu Asn Asn Ala Leu Asp Met Ala Pro Glu Ile
  1                 5                  10                  15
His Met Thr Gly Pro Met Cys Leu Ile Glu Asn Thr Asn Gly Glu
                 20                  25                  30
Leu Val Ala Asn Pro Glu Ala Leu Lys Ile Leu Ser Ala Ile Thr
                 35                  40                  45
Gln Pro Val Val Val Val Ala Ile Val Gly Leu Tyr Arg Thr Gly
                 50                  55                  60
Lys Ser Tyr Leu Met Asn Lys Leu Ala Gly Lys Asn Lys Gly Phe
                 65                  70                  75
Ser Leu Gly Ser Thr Val Lys Ser His Thr Lys Gly Ile Trp Met
                 80                  85                  90
Trp Cys Val Pro His Pro Lys Lys Pro Glu His Thr Leu Val Leu
                 95                 100                 105
Leu Asp Thr Glu Gly Leu Gly Asp Val Lys Lys Gly Asp Asn Gln
                110                 115                 120
Asn Asp Ser Trp Ile Phe Thr Leu Ala Val Leu Leu Ser Ser Thr
                125                 130                 135
Leu Val Tyr Asn Ser Met Gly Thr Ile Asn Gln Gln Ala Met Asp
                140                 145                 150
Gln Leu Tyr Tyr Val Thr Glu Leu Thr His Arg Ile Arg Ser Lys
                155                 160                 165
Ser Ser Pro Asp Glu Asn Glu Asn Glu Asp Ser Ala Asp Phe Val
                170                 175                 180
Ser Phe Phe Pro Asp Phe Val Trp Thr Leu Arg Asp Phe Ser Leu
                185                 190                 195
Asp Leu Glu Ala Asp Gly Gln Pro Leu Thr Pro Asp Glu Tyr Leu
                200                 205                 210
Glu Tyr Ser Leu Lys Leu Thr Gln Gly Thr Ser Gln Lys Asp Lys
                215                 220                 225
Asn Phe Asn Leu Pro Gln Leu Cys Ile Trp Lys Phe Phe Pro Lys
                230                 235                 240
Lys Lys Cys Phe Val Phe Asp Leu Pro Ile His Arg Arg Lys Leu
                245                 250                 255
Ala Gln Leu Glu Lys Leu Gln Asp Glu Glu Leu Asp Pro Glu Phe
                260                 265                 270
Val Gln Gln Val Ala Asp Phe Cys Ser Tyr Ile Phe Ser Asn Ser
                275                 280                 285
Lys Thr Lys Thr Leu Ser Gly Gly Ile Lys Val Asn Gly Pro Arg
                290                 295                 300
Leu Glu Ser Leu Val Leu Thr Tyr Ile Asn Ala Ile Ser Arg Gly
                305                 310                 315
```

```
Asp Leu Pro Cys Met Glu Asn Ala Val Leu Ala Leu Ala Gln Ile
                320                 325                 330
Glu Asn Ser Ala Ala Val Gln Lys Ala Ile Ala His Tyr Asp Gln
                335                 340                 345
Gln Met Gly Gln Lys Val Gln Leu Pro Ala Glu Thr Leu Gln Glu
                350                 355                 360
Leu Leu Asp Leu His Arg Val Ser Glu Arg Glu Ala Thr Glu Val
                365                 370                 375
Tyr Met Lys Asn Ser Phe Lys Asp Val Asp His Leu Phe Gln Lys
                380                 385                 390
Lys Leu Ala Ala Gln Leu Asp Lys Lys Arg Asp Asp Phe Cys Lys
                395                 400                 405
Gln Asn Gln Glu Ala Ser Ser Asp Arg Cys Ser Ala Leu Leu Gln
                410                 415                 420
Val Ile Phe Ser Pro Leu Glu Glu Glu Val Lys Ala Gly Ile Tyr
                425                 430                 435
Ser Lys Pro Gly Gly Tyr Cys Leu Phe Ile Gln Lys Leu Gln Asp
                440                 445                 450
Leu Glu Lys Lys Tyr Tyr Glu Glu Pro Arg Lys Gly Ile Gln Ala
                455                 460                 465
Glu Glu Ile Leu Gln Thr Tyr Leu Lys Ser Lys Glu Ser Val Thr
                470                 475                 480
Asp Ala Ile Leu Gln Thr Asp Gln Ile Leu Thr Glu Lys Glu Lys
                485                 490                 495
Glu Ile Glu Val Glu Cys Val Lys Ala Glu Ser Ala Gln Ala Ser
                500                 505                 510
Ala Lys Met Val Glu Glu Met Gln Ile Lys Tyr Gln Gln Met Met
                515                 520                 525
Glu Glu Lys Glu Lys Ser Tyr Gln Glu His Val Lys Gln Leu Thr
                530                 535                 540
Glu Lys Met Glu Arg Glu Arg Ala Gln Leu Leu Glu Glu Gln Glu
                545                 550                 555
Lys Thr Leu Thr Ser Lys Leu Gln Val Ser Lys Cys Lys Xaa Xaa
                560                 565                 570
```

<210> 64
<211> 168
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:402231.6.orf3:2000MAY19

<400> 64

```
Ala Leu Phe Ser Arg Ile Ile Gln Gln Leu Val Asn Gly Ile Ile
  1                 5                  10                  15
Thr Pro Ala Thr Ile Pro Ser Leu Gly Pro Trp Gly Val Leu His
                 20                  25                  30
Ser Asn Pro Met Asp Tyr Ala Trp Gly Ala Asn Gly Leu Asp Ala
                 35                  40                  45
Ile Ile Thr Gln Leu Leu Asn Gln Phe Glu Asn Thr Gly Pro Pro
                 50                  55                  60
Pro Ala Asp Lys Glu Lys Ile Gln Ala Leu Pro Thr Val Pro Val
                 65                  70                  75
Thr Glu Glu His Val Gly Ser Gly Leu Glu Cys Pro Val Cys Lys
                 80                  85                  90
Asp Asp Tyr Ala Leu Gly Glu Arg Val Arg Gln Leu Pro Cys Asn
                 95                 100                 105
His Leu Phe His Thr Thr Tyr Glu Gln Ala Trp Leu Glu Gln His
                110                 115                 120
Asp Ser Cys Pro Val Cys Arg Lys Ser Leu Thr Gly Gln Asn Thr
                125                 130                 135
Ala Thr Asn Pro Pro Gly Leu Thr Gly Val Ser Phe Ser Ser Ser
                140                 145                 150
Ser Ser Ser Ser Ser Ser Ser Ser Pro Ser Asn Glu Asn Ala Thr
```

155 160 165
Ser Asn Ser

<210> 65
<211> 246
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1076157.1.orf3:2000MAY19

<220>
<221> unsure
<222> 240
<223> unknown or other

<400> 65
Pro Lys Gln Gly Ile Asn Val Trp Ser Pro Arg His Pro Glu Asn
1 5 10 15
Phe Leu Gly Ile Glu Ser Arg Pro Pro Met Leu Ser Leu Ser Pro
20 25 30
Ile Leu Leu Tyr Thr Cys Glu Met Phe Gln Asp Pro Val Ala Phe
35 40 45
Lys Asp Val Ala Val Asn Phe Thr Gln Glu Glu Trp Ala Leu Leu
50 55 60
Asp Ile Ser Gln Arg Lys Leu Tyr Arg Glu Val Met Leu Glu Thr
65 70 75
Phe Arg Asn Leu Thr Ser Ile Gly Lys Lys Trp Lys Asp Gln Asn
80 85 90
Ile Glu Tyr Glu Tyr Gln Asn Pro Arg Arg Asn Phe Arg Ser Leu
95 100 105
Ile Glu Gly Asn Val Asn Glu Ile Lys Glu Asp Ser His Cys Gly
110 115 120
Glu Thr Phe Thr Gln Val Pro Asp Asp Arg Leu Asn Phe Gln Glu
125 130 135
Lys Lys Ala Ser Pro Glu Ala Lys Ser Cys Asp Asn Phe Val Cys
140 145 150
Gly Glu Val Gly Ile Gly Asn Ser Ser Phe Asn Met Asn Ile Arg
155 160 165
Gly Asp Ile Gly His Lys Ala Tyr Glu Tyr Gln Asp Tyr Ala Pro
170 175 180
Lys Pro Tyr Lys Cys Gln Gln Pro Lys Lys Ala Phe Arg Tyr His
185 190 195
Pro Ser Phe Arg Thr Gln Glu Arg Asn His Thr Gly Glu Lys Pro
200 205 210
Tyr Ala Cys Lys Glu Cys Gly Lys Thr Phe Ile Ser His Ser Gly
215 220 225
Ile Arg Arg Arg Met Val Met His Ser Gly Asp Gly Pro Leu Xaa
230 235 240
Val Ser Phe Val Gly Lys
245

<210> 66
<211> 120
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1083142.1.orf3:2000MAY19

<220>
<221> unsure
<222> 1
<223> unknown or other

<400> 66
Xaa Phe Pro Val Leu Glu Pro His Gln Val Gly Leu Ile Arg Ser
1 5 10 15
Tyr Asn Ser Lys Thr Met Thr Cys Phe Gln Glu Leu Val Thr Phe
20 25 30
Arg Asp Val Ala Ile Asp Phe Ser Arg Gln Glu Trp Glu Tyr Leu
35 40 45
Asp Pro Asn Gln Arg Asp Leu Tyr Arg Asp Val Met Leu Glu Asn
50 55 60
Tyr Arg Asn Leu Val Ser Leu Gly Gly His Ser Ile Ser Lys Pro
65 70 75
Val Val Val Asp Leu Leu Glu Arg Gly Lys Glu Pro Trp Met Ile
80 85 90
Leu Arg Glu Glu Thr Gln Phe Thr Asp Leu Asp Leu Gln Cys Glu
95 100 105
Ile Ile Ser Tyr Ile Glu Val Pro Thr Tyr Glu Thr Asp Ile Ser
110 115 120

<210> 67
<211> 122
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:1083264.1.orf2:2000MAY19

<400> 67
Lys Lys Ser Gln Lys Glu Ser Thr Gln Gln Thr Arg Ile His Phe
1 5 10 15
Gln Arg Asp Ile Leu Cys Lys Glu Ala Thr Trp Lys Arg Lys Glu
20 25 30
Lys Lys Ser Gly Met Ala Leu Thr Gln Gly Pro Leu Lys Phe Met
35 40 45
Asp Val Ala Ile Glu Phe Ser Gln Glu Glu Trp Lys Cys Leu Asp
50 55 60
Pro Ala Gln Arg Thr Leu Tyr Arg Asp Val Met Leu Glu Asn Tyr
65 70 75
Arg Asn Leu Val Ser Leu Gly Ile Cys Leu Pro Asp Leu Ser Val
80 85 90
Thr Ser Met Leu Glu Gln Lys Arg Asp Pro Trp Thr Leu Gln Ser
95 100 105
Glu Glu Lys Ile Ala Asn Asp Pro Asp Gly Arg Glu Cys Ile Gln
110 115 120
Lys Val

<210> 68
<211> 428
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:350793.2.orf3:2000MAY19

<400> 68
Ala Gln Gly Ser Ser Trp Lys Leu Pro Phe Glu Arg Leu Ala Phe
1 5 10 15
Val Leu Ser Ser Asn Ser Leu Lys His Cys Thr Glu Leu Glu Leu
20 25 30
Phe Lys Ala Thr Cys Arg Trp Leu Arg Leu Glu Glu Pro Arg Met
35 40 45
Asp Phe Ala Ala Lys Leu Met Lys Asn Ile Arg Phe Pro Leu Met
50 55 60
Thr Pro Gln Glu Leu Ile Asn Tyr Val Gln Thr Val Asp Phe Met

```
                   65                  70                  75
Arg Thr Asp Asn Thr Cys Val Asn Leu Leu Leu Glu Ala Ser Asn
                   80                  85                  90
Tyr Gln Met Met Pro Tyr Met Gln Pro Val Met Gln Ser Asp Arg
                   95                 100                 105
Thr Ala Ile Arg Ser Asp Thr Thr His Leu Val Thr Leu Gly Gly
                  110                 115                 120
Val Leu Arg Gln Gln Leu Val Val Ser Lys Glu Leu Arg Met Tyr
                  125                 130                 135
Asp Glu Lys Ala His Glu Trp Lys Ser Leu Ala Pro Met Asp Ala
                  140                 145                 150
Pro Arg Tyr Gln His Gly Ile Ala Val Ile Gly Asn Phe Leu Tyr
                  155                 160                 165
Val Val Gly Gly Gln Ser Asn Tyr Asp Thr Lys Gly Lys Thr Ala
                  170                 175                 180
Val Asp Thr Val Phe Arg Phe Asp Pro Arg Tyr Asn Lys Trp Met
                  185                 190                 195
Gln Val Ala Ser Leu Asn Glu Lys Arg Thr Phe Phe His Leu Ser
                  200                 205                 210
Ala Leu Lys Gly Tyr Leu Tyr Ala Val Gly Gly Arg Asn Ala Ala
                  215                 220                 225
Gly Glu Leu Pro Thr Val Glu Cys Tyr Asn Pro Arg Thr Asn Glu
                  230                 235                 240
Trp Thr Tyr Val Ala Lys Met Ser Glu Pro His Tyr Gly His Ala
                  245                 250                 255
Gly Thr Val Tyr Gly Gly Val Met Tyr Ile Ser Gly Gly Ile Thr
                  260                 265                 270
His Asp Thr Phe Gln Lys Glu Leu Met Cys Phe Asp Pro Asp Thr
                  275                 280                 285
Asp Lys Trp Ile Gln Lys Ala Pro Met Thr Thr Val Arg Gly Leu
                  290                 295                 300
His Cys Met Cys Thr Val Gly Glu Arg Leu Tyr Val Ile Gly Gly
                  305                 310                 315
Asn His Phe Arg Gly Thr Ser Asp Tyr Asp Asp Val Leu Ser Cys
                  320                 325                 330
Glu Tyr Tyr Ser Pro Ile Leu Asp Gln Trp Thr Pro Ile Ala Ala
                  335                 340                 345
Met Leu Arg Gly Gln Ser Asp Val Gly Val Ala Val Phe Glu Asn
                  350                 355                 360
Lys Ile Tyr Val Val Gly Gly Tyr Ser Trp Asn Asn Arg Cys Met
                  365                 370                 375
Val Glu Ile Val Gln Lys Tyr Asp Pro Asp Lys Asp Glu Trp His
                  380                 385                 390
Lys Val Phe Asp Leu Pro Glu Ser Leu Gly Gly Ile Arg Ala Cys
                  395                 400                 405
Thr Leu Thr Val Phe Pro Pro Glu Glu Thr Thr Pro Ser Pro Ser
                  410                 415                 420
Arg Glu Ser Pro Leu Ser Ala Pro
                  425
```

<210> 69
<211> 307
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LG:408751.3.orf2:2000MAY19

<400> 69

```
Arg Asp Pro Gly Trp Gln Ile Arg Asp Arg Ala Gly Leu Ala Trp
  1                 5                  10                  15
Asn Met Leu Ala Asn Ser Ala Ser Val Arg Ile Leu Ile Lys Gly
                  20                  25                  30
Gly Lys Val Val Asn Asp Asp Cys Thr His Glu Ala Asp Val Tyr
                  35                  40                  45
Ile Glu Asn Gly Ile Ile Gln Gln Val Gly Arg Glu Leu Met Ile
```

```
                 50                  55                  60
Pro Gly Gly Ala Lys Val Ile Asp Ala Thr Gly Lys Leu Val Ile
                 65                  70                  75
Pro Gly Gly Ile Asp Thr Ser Thr His Phe His Gln Thr Phe Met
                 80                  85                  90
Asn Ala Thr Cys Val Asp Asp Phe Tyr His Gly Thr Lys Ala Ala
                 95                 100                 105
Leu Val Gly Gly Thr Thr Met Ile Ile Gly His Val Leu Pro Asp
                110                 115                 120
Lys Glu Thr Ser Leu Val Asp Ala Tyr Glu Lys Cys Arg Gly Leu
                125                 130                 135
Ala Asp Pro Lys Val Cys Cys Asp Tyr Ala Leu His Val Gly Ile
                140                 145                 150
Thr Trp Trp Ala Pro Lys Val Lys Ala Glu Met Glu Thr Leu Val
                155                 160                 165
Arg Glu Lys Gly Val Asn Ser Phe Gln Met Phe Met Thr Tyr Lys
                170                 175                 180
Asp Leu Tyr Met Leu Arg Asp Ser Glu Leu Tyr Gln Val Leu His
                185                 190                 195
Ala Cys Lys Asp Ile Gly Ala Ile Ala Arg Val His Ala Glu Asn
                200                 205                 210
Gly Glu Leu Val Ala Glu Gly Ala Lys Glu Ala Leu Asp Leu Gly
                215                 220                 225
Ile Thr Gly Pro Glu Gly Ile Glu Ile Ser Arg Pro Glu Glu Leu
                230                 235                 240
Glu Ala Glu Ala Thr His Arg Val Ile Thr Arg Asp Gly Gly Asn
                245                 250                 255
His Asp Ala Ala Ser Trp Cys Ser Ala His His Leu Tyr Pro Cys
                260                 265                 270
Gln Pro Ser Leu Gly His Gly Pro Trp Ala Asp Val Lys Glu Pro
                275                 280                 285
Ser Ser Ser Gly Gly Gly Gln Leu Gly Arg Ala Ser Leu Leu Gly
                290                 295                 300
Leu Gly Lys Leu Tyr Leu Leu
                305
```

<210> 70
<211> 198
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:336120.1.orf1:2000MAY01

<400> 70

```
Ile Ile Pro Gln Arg Ser Asn Gly Asp Arg Trp Gly Arg Ser Leu
  1               5                  10                  15
Leu Pro Ser Arg Thr Phe Leu Gln Ala Leu Asn Leu Gly Ile Glu
                 20                  25                  30
Val Ile Asn Thr Thr Asp Tyr Leu His Phe Ser Lys Glu Cys Ser
                 35                  40                  45
Arg Ala Leu Leu Lys Met Gln Tyr Cys Pro His Cys Gln Gly Leu
                 50                  55                  60
Ala Leu Thr Lys Pro Cys Met Gly Tyr Cys Leu Asn Val Met Arg
                 65                  70                  75
Gly Cys Leu Ala His Met Ala Glu Leu Asn Pro His Trp His Ala
                 80                  85                  90
Tyr Ile Arg Ser Leu Glu Glu Leu Ser Asp Ala Met His Gly Thr
                 95                 100                 105
Tyr Asp Ile Gly His Val Leu Leu Asn Phe His Leu Leu Val Asn
                110                 115                 120
Asp Ala Val Leu Gln Ala His Leu Asn Gly Gln Lys Leu Leu Glu
                125                 130                 135
Gln Val Asn Arg Ile Cys Gly Arg Pro Val Arg Thr Pro Thr Gln
                140                 145                 150
Ser Pro Arg Cys Ser Phe Asp Gln Ser Lys Glu Lys His Gly Met
```

```
                155                 160                 165
Lys Thr Thr Thr Arg Asn Ser Glu Glu Thr Leu Ala Asn Arg Arg
                170                 175                 180
Lys Glu Phe Ile Asn Ser Leu Ser Thr Val Gln Val Ile Leu Trp
                185                 190                 195
Arg Ser Ser
```

<210> 71
<211> 227
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:234104.2.orf1:2000MAY01

<400> 71

```
Ala Thr Pro Ser Gly Arg Pro Gln Ser Trp Thr Arg Phe Ser Leu
  1               5                  10                  15
Trp Arg Gly Pro Arg Arg Thr Arg Pro Ser Pro Pro Ala Pro Ala
                20                  25                  30
Pro Ala Gly Met Gly Ser Glu His Asp Gly Arg Ser Gly Pro Val
                35                  40                  45
Leu Thr Pro Ala Asp Thr Leu His Pro Pro Thr Arg Leu Gln Pro
                50                  55                  60
Ser Pro Pro Asp Thr His Pro Gly Gly Ser Ser Leu Pro Ala Pro
                65                  70                  75
Arg Pro Ala Leu Ser Cys Trp Ala Arg Val Phe Ala Ser Leu Val
                80                  85                  90
Arg Pro Ala Gly Phe Pro Gly Gly Thr His Gly Ala Pro Gly Met
                95                 100                 105
Pro Leu Gly Ser Pro Ser Thr Ser Thr Ala Gln Trp Pro Tyr Val
               110                 115                 120
Gln Leu Val Pro Gly Pro Arg Val Arg Lys Thr Ala Ser Arg Ser
               125                 130                 135
His Cys Gln Glu Arg Ala Glu Glu Trp Ser Gly Pro Arg Arg Pro
               140                 145                 150
Trp Gly Glu Gly Asp Pro Gly Pro Val Thr Ala Thr Pro Gly Thr
               155                 160                 165
Pro Gly Gly Ala Pro Thr Ser Ala Phe Ser Cys Ala Ala Lys Leu
               170                 175                 180
Gln Lys Pro Asp Ala Gly Leu Val Val Ala Asn Gly Thr Met Cys
               185                 190                 195
Cys Pro Ala Lys His Thr Trp Arg Ser Gly Pro Lys Ile Pro Ile
               200                 205                 210
Leu Asp Phe His Pro Ala Pro Ser Ser Thr Pro Arg Ser Ala Leu
               215                 220                 225
Ser His
```

<210> 72
<211> 122
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:450887.1.orf3:2000MAY01

<400> 72

```
Ser Val His Phe Ser Arg Lys Gly Phe Val Leu Met Ala Pro Pro
  1               5                  10                  15
Gln Pro Lys Ser Gly Leu Phe Val Gly Ile Asn Lys Gly His Val
                20                  25                  30
Val Thr Lys Arg Glu Leu Pro Pro Arg Pro Cys His Arg Lys Gly
                35                  40                  45
```

```
Lys Ser Thr Lys Arg Val Ser Met Val Arg Gly Leu Ile Arg Glu
                50                  55                  60
Val Ala Gly Phe Ala Pro Tyr Glu Lys Arg Ile Thr Glu Leu Leu
                65                  70                  75
Lys Val Gly Lys Asp Lys Arg Ala Leu Lys Leu Ala Lys Arg Lys
                80                  85                  90
Leu Gly Thr His Lys Arg Ala Lys Lys Lys Arg Glu Glu Met Ala
                95                 100                 105
Gly Val Leu Arg Lys Met Arg Ser Ala Gly Thr His Thr Asp Lys
               110                 115                 120
Lys Lys
```

<210> 73
<211> 209
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:119992.3.orf2:2000MAY01

<400> 73

```
Cys Ser Gln Ile Glu Leu Ala Ile Glu Leu Asp Ser Thr His Leu
  1                 5                  10                  15
Val Thr Leu Gly Gly Val Leu Arg Gln Gln Leu Val Val Ser Lys
                20                  25                  30
Glu Leu Arg Met Tyr Asp Glu Arg Ala Gln Glu Trp Arg Ser Leu
                35                  40                  45
Ala Pro Met Asp Ala Pro Arg Tyr Gln His Gly Tyr Trp Leu Phe
                50                  55                  60
Ile Gly Asn Phe Leu Tyr Val Val Gly Gly Gln Ser Asn Tyr Asp
                65                  70                  75
Thr Lys Gly Lys Thr Ala Val Asp Thr Val Phe Arg Phe Asp Pro
                80                  85                  90
Arg Tyr Asn Lys Trp Met Gln Val Ala Ser Leu Asn Glu Lys Arg
                95                 100                 105
Thr Phe Phe His Leu Ser Ala Leu Lys Gly His Leu Tyr Ala Val
               110                 115                 120
Gly Gly Arg Ser Ala Ala Gly Glu Leu Gly Thr Val Glu Cys Tyr
               125                 130                 135
Asn Pro Arg Met Asn Glu Trp Ser Tyr Val Ala Lys Met Ser Glu
               140                 145                 150
Pro His Tyr Gly His Ala Gly Thr Val Tyr Gly Gly Leu Met Tyr
               155                 160                 165
Ile Ser Gly Gly Ile Thr His Asp Thr Phe Gln Asn Glu Leu Met
               170                 175                 180
Cys Phe Asp Pro Asp Thr Asp Lys Trp Met Gln Lys Ala Pro Met
               185                 190                 195
Thr Thr Val Arg Gly Leu His Cys Met Cys Thr Arg Trp Arg
               200                 205
```

<210> 74
<211> 312
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:197241.2.orf1:2000MAY01

<400> 74

```
Tyr Ser Arg Ile Leu Ile Leu Gln Met Phe Ile Leu Gly Ala Ile
  1                 5                  10                  15
Ile Gln Ile Leu Pro Trp Val Met Ala Ser Gln Asn Ser Lys His
                20                  25                  30
His Pro Glu Leu Val Asp Leu Phe Ser Arg Ser Gly Ile Tyr Ile
```

```
                35                  40                  45
Lys Gln Val Val Leu Cys Lys Phe His Ser Val Phe Leu Ser Gln
                50                  55                  60
Lys Gly Gln Val Tyr Thr Cys Gly His Gly Pro Gly Arg Ala Ile
                65                  70                  75
Arg Asp Met Gly Asp Glu Gln Thr Cys Leu Val Pro Arg Leu Val
                80                  85                  90
Glu Gly Leu Asn Gly His Asn Cys Ser Gln Val Ala Ala Ala Lys
                95                  100                 105
Asp His Thr Val Val Leu Thr Glu Asp Gly Cys Val Tyr Thr Phe
                110                 115                 120
Gly Leu Asn Ile Phe His Gln Leu Gly Ile Ile Pro Pro Pro Ser
                125                 130                 135
Ser Cys Asn Val Pro Arg Gln Ile Gln Ala Lys Tyr Leu Lys Gly
                140                 145                 150
Arg Thr Ile Ile Gly Val Ala Ala Gly Arg Phe His Thr Val Leu
                155                 160                 165
Trp Thr Arg Glu Ala Val Tyr Thr Met Gly Leu His Gly Gly Gln
                170                 175                 180
Leu Gly Cys Leu Leu Asp Pro Asn Gly Glu Lys Cys Val Thr Ala
                185                 190                 195
Pro Arg Gln Val Ser Ala Leu His His Lys Asp Ile Ala Leu Ser
                200                 205                 210
Leu Val Ala Ala Ser Asp Gly Ala Thr Val Cys Val Thr Thr Arg
                215                 220                 225
Gly Asp Ile Tyr Leu Leu Ala Asp Tyr Gln Cys Lys Lys Met Ala
                230                 235                 240
Ser Lys Gln Leu Asn Leu Lys Lys Val Leu Val Ser Gly Gly His
                245                 250                 255
Met Glu Tyr Lys Val Asp Pro Glu His Leu Lys Glu Asn Gly Gly
                260                 265                 270
Gln Lys Ile Cys Ile Leu Ala Met Asp Gly Ala Gly Arg Val Phe
                275                 280                 285
Cys Trp Arg Ser Val Asn Ser Ser Leu Lys Gln Cys Arg Leu Gly
                290                 295                 300
Leu Ser Thr Ser Gly Ser Ser Phe Leu Ile Trp Leu
                305                 310
```

<210> 75
<211> 190
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:406860.20.orf3:2000MAY01

<400> 75

```
Leu Tyr Val Met Leu Glu Met Thr Arg Pro Ser Ser Leu Ser Leu
  1               5                  10                  15
Ser Gln Leu Ala Leu Phe Ser Arg Ala Val Leu Pro Val Gly Arg
                20                  25                  30
Ala Glu Asp Leu Ala Gly Glu Ala Gly Glu Ala Cys Trp Pro Ser
                35                  40                  45
Leu Cys Ala Pro Leu His Ala His Pro Pro Ala Pro Pro Glu Arg
                50                  55                  60
Ile Val His Pro Ala Ala Arg Ser Leu Asp Leu His Phe Gly Ala
                65                  70                  75
Pro Gly Arg Val Glu Leu Arg Cys Glu Val Ala Pro Ala Gly Ser
                80                  85                  90
Gln Val Arg Trp Tyr Lys Asp Gly Leu Glu Val Glu Ala Ser Asp
                95                  100                 105
Ala Leu Gln Leu Gly Ala Glu Gly Pro Thr Arg Thr Leu Thr Leu
                110                 115                 120
Pro His Ala Gln Pro Glu Asp Ala Gly Glu Tyr Val Cys Glu Thr
                125                 130                 135
Arg His Glu Ala Ile Thr Phe Asn Val Ile Leu Ala Glu Pro Pro
```

```
                  140                 145                 150
Val Gln Phe Leu Ala Leu Glu Thr Thr Pro Ser Pro Leu Cys Val
                  155                 160                 165
Gly Pro Gly Glu Pro Val Val Gln Glu Gly Glu Gly Leu Glu Leu
                  170                 175                 180
His Ala Glu Gly Pro Ala Glu Ser Leu His
                  185                 190

<210> 76
<211> 295
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:142384.1.orf3:2000MAY01

<400> 76
Arg Thr Cys Cys Arg Val Val Pro Glu Ala Lys Gln Arg Trp Arg
  1                 5                  10                  15
Arg Val Arg Leu Arg Arg Arg Gln Arg Arg Ala Pro Gly Arg Arg
                 20                  25                  30
Ala Pro Gly Arg Ala Ala Leu Leu Val Leu Leu Ala Leu Ala Ala
                 35                  40                  45
Ala Ala Ala Gly Ser Gly Arg Leu Ser Cys Arg Met Cys Gly Arg
                 50                  55                  60
Arg Arg Arg Ser Val Gly Gly Ala Gly Gly Pro Gly Ser Gly Leu
                 65                  70                  75
Ala Pro Leu Pro Gly Leu Pro Pro Ser Ala Ala Ala His Gly Ala
                 80                  85                  90
Ala Leu Leu Ser His Trp Asp Pro Thr Leu Ser Ser Asp Trp Asp
                 95                 100                 105
Gly Glu Arg Thr Ala Pro Gln Cys Leu Leu Arg Ile Lys Arg Asp
                110                 115                 120
Ile Met Ser Ile Tyr Lys Glu Pro Pro Pro Gly Met Phe Val Val
                125                 130                 135
Pro Asp Thr Val Asp Met Thr Lys Ile His Ala Leu Ile Thr Gly
                140                 145                 150
Pro Phe Asp Thr Pro Tyr Glu Gly Gly Phe Phe Leu Phe Val Phe
                155                 160                 165
Arg Cys Pro Pro Asp Tyr Pro Ile His Pro Pro Arg Val Lys Leu
                170                 175                 180
Met Thr Thr Gly Asn Asn Thr Val Arg Phe Asn Pro Asn Phe Tyr
                185                 190                 195
Arg Asn Gly Lys Val Cys Leu Ser Ile Leu Gly Thr Trp Thr Gly
                200                 205                 210
Pro Ala Trp Ser Pro Ala Gln Ser Ile Ser Ser Val Leu Ile Ser
                215                 220                 225
Ile Gln Ser Leu Met Thr Glu Asn Pro Tyr His Asn Glu Pro Gly
                230                 235                 240
Phe Glu Gln Glu Arg His Pro Gly Asp Ser Lys Asn Tyr Asn Glu
                245                 250                 255
Cys Ile Arg His Glu Thr Ile Arg Val Ala Val Cys Asp Met Met
                260                 265                 270
Glu Gly Lys Cys Pro Cys Pro Glu Pro Leu Arg Gly Val Met Glu
                275                 280                 285
Lys Ser Phe Leu Glu Tyr Tyr Asp Phe Tyr
                290                 295

<210> 77
<211> 288
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:895427.1.orf2:2000MAY01
```

<400> 77
Ala Pro Arg Leu Trp Ala Cys Pro Cys His Cys Trp Trp Ser Gly
1 5 10 15
Ser Gly Pro Pro Ala Arg Cys Pro Tyr Ile Ile Gln Lys Cys Val
20 25 30
Gly Gln Ile Glu Arg Arg Gly Leu Arg Val Val Gly Leu Tyr Arg
35 40 45
Leu Cys Gly Ser Ala Ala Val Lys Lys Glu Leu Arg Asp Ala Phe
50 55 60
Glu Arg Asp Ser Ala Ala Val Cys Leu Ser Glu Asp Leu Tyr Pro
65 70 75
Asp Ile Asn Val Ile Thr Gly Ile Leu Lys Asp Tyr Leu Arg Glu
80 85 90
Leu Pro Thr Pro Leu Ile Thr Gln Pro Leu Tyr Lys Val Val Leu
95 100 105
Glu Ala Met Ala Pro Gly Thr Pro Gln Thr Glu Phe Pro Pro Pro
110 115 120
Leu Arg Ala Pro Glu Gly Ser Tyr Ser Cys Leu Pro Asp Val Glu
125 130 135
Arg Ala Thr Leu Thr Leu Leu Leu Asp His Leu Arg Leu Val Ser
140 145 150
Ser Phe His Ala Tyr Asn Arg Met Thr Pro Gln Asn Leu Ala Val
155 160 165
Cys Phe Gly Pro Val Leu Leu Pro Ala Arg Gln Ala Pro Thr Arg
170 175 180
Pro Arg Ala Arg Ser Ser Gly Pro Gly Leu Ala Ser Ala Val Asp
185 190 195
Phe Lys His His Ile Glu Val Leu His Tyr Leu Leu Gln Ser Trp
200 205 210
Pro Asp Pro Arg Leu Pro Arg Gln Ser Pro Asp Val Ala Pro Tyr
215 220 225
Leu Arg Pro Lys Arg Gln Pro Pro Leu His Leu Pro Leu Ala Asp
230 235 240
Pro Glu Val Val Thr Arg Pro Arg Gly Arg Gly Gly Pro Glu Ser
245 250 255
Pro Pro Ser Asn Arg Tyr Ala Gly Asp Trp Ser Val Cys Gly Arg
260 265 270
Gly Leu Pro Asp Leu Trp Ala Gly Phe Pro Val Arg Ala Arg Leu
275 280 285
Arg Pro Leu

<210> 78
<211> 294
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:757439.1.orf1:2000MAY01

<400> 78
Leu Ala Ala Pro Gln Ser His Ser Ile Pro Ser Pro Pro Gly Ala
1 5 10 15
His Leu Leu Lys Thr Arg Val Leu Pro Ser Ala Arg Arg Ala Arg
20 25 30
Ala Arg Gly Ala Arg Glu Leu Arg Ser Ala Arg Ala Met Gly Pro
35 40 45
Pro Pro Gly Ala Gly Val Ser Cys Arg Gly Gly Cys Gly Phe Ser
50 55 60
Arg Leu Leu Ala Trp Cys Phe Leu Leu Ala Leu Ser Pro Gln Ala
65 70 75
Pro Gly Ser Arg Gly Ala Glu Ala Val Trp Thr Ala Tyr Leu Asn
80 85 90
Val Ser Trp Arg Val Pro His Thr Gly Val Asn Arg Thr Val Trp
95 100 105
Glu Leu Ser Glu Glu Gly Val Tyr Gly Pro Asp Ser Pro Leu Glu

```
                110                 115                 120
Pro Val Ala Gly Val Leu Val Pro Pro Asp Gly Pro Gly Ala Leu
                125                 130                 135
Asn Ala Cys Asn Pro His Thr Asn Phe Thr Val Pro Thr Val Trp
                140                 145                 150
Gly Ser Thr Val Gln Val Ser Trp Leu Gly Leu Ile Gln Arg Gly
                155                 160                 165
Gly Gly Cys Thr Phe Ala Asp Lys Ile His Leu Ala Tyr Glu Arg
                170                 175                 180
Gly Ala Ser Gly Ala Val Ile Phe Asn Phe Pro Gly Thr Arg Asn
                185                 190                 195
Glu Val Ile Pro Met Ser His Pro Gly Ala Val Asp Ile Val Ala
                200                 205                 210
Ile Met Ile Arg Gln Ser Glu Arg His Lys Asn Ser Ala Ile Tyr
                215                 220                 225
Ser Lys Arg His Thr Ser Asp Asn Gly His Arg Ser Arg Glu Lys
                230                 235                 240
Thr Trp Pro Leu Gly Glu Ser Leu Phe Asn Phe Phe Arg Phe Leu
                245                 250                 255
Cys Pro Phe Leu Leu Leu Arg Arg Ala Thr Val Gly Tyr Phe Ile
                260                 265                 270
Phe Tyr Ser Ala Arg Arg Leu Arg Asn Ala Arg Ala Gln Ser Arg
                275                 280                 285
Lys Gln Arg Pro Ile Lys Gly Arg Cys
                290
```

<210> 79
<211> 196
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:1144066.1.orf3:2000MAY01

<400> 79

```
Gly Ala Thr Pro Arg Ala Gly Glu Arg Ala Pro Leu Leu Pro Asp
  1                 5                  10                  15
Arg Ala Ala His Ala Ala Ser Gly Thr Ile Thr Val Ala Gly Arg
                 20                  25                  30
Arg Pro Val Gln Ile Leu Ser Glu Phe Phe Gly Ala Phe Ser Pro
                 35                  40                  45
Arg Lys Leu Ala Ile Gln Lys Cys Ala Ser Arg Thr Ala Ala Ala
                 50                  55                  60
Met Gly Ser Glu Asp His Gly Ala Gln Lys Pro Ser Cys Lys Ile
                 65                  70                  75
Met Thr Phe Arg Pro Thr Met Gly Glu Phe Lys Asp Phe Asn Lys
                 80                  85                  90
Tyr Val Gly Tyr Ile Glu Ser Gln Gly Ala His Arg Ala Gly Leu
                 95                 100                 105
Gly Lys Ile Ile Pro Pro Lys Glu Trp Lys Pro Arg Gln Thr Tyr
                110                 115                 120
Asp Asp Ile Asp Asp Val Val Ile Pro Gly Pro Ile Gln Gln Val
                125                 130                 135
Val Thr Gly Gln Ser Gly Leu Phe Thr Gln Tyr Asn Ile Gln Lys
                140                 145                 150
Lys Gly Met Thr Val Gly Glu Tyr Arg Arg Leu Gly Asn Ser Glu
                155                 160                 165
Lys Tyr Cys Thr Pro Arg Asp Gln Asp Phe Asp Asp Leu Glu Arg
                170                 175                 180
Lys Tyr Trp Glu Gly Thr Leu Thr Leu Cys Leu Pro Asp Leu Arg
                185                 190                 195
Gly
```

<210> 80
<211> 745

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:243660.4.orf3:2000MAY01

<400> 80
Glu Gly Trp Thr Gln Pro Gln Gln Ala Gly Glu Gly Pro His Pro
1 5 10 15
Ala Ala His Glu Cys Leu His Asp Leu Gln Gln Ala Ala Pro Gly
20 25 30
Pro Gly Pro Pro Ala Ser Ser Gln Pro Gly Gln Pro Asp Arg Gln
35 40 45
Gln Asp Pro Gly Arg Val Val Val Cys Pro Gly Ala Gln Gly Glu
50 55 60
Ala Glu Val Pro Arg Pro Gly Leu Pro Gly Glu Gly Gly Pro Leu
65 70 75
Gln Gly Pro Pro Ser Ile Gly Ser Gly Ala Thr Arg Thr Glu Arg
80 85 90
Ser Pro Ala Gln Arg Pro Ser Pro Arg Ser Leu Gly Leu Ala Gly
95 100 105
Gly His Lys Glu Thr Arg Glu Arg Ser Met Ser Glu Thr Gly Thr
110 115 120
Ala Ala Cys Pro Trp Val Cys Pro Arg Glu Leu Leu Ser Val Ala
125 130 135
Ala Gln Thr Leu Leu Ser Ser Asp Thr Lys Ala Pro Gly Ser Ser
140 145 150
Ser Cys Gly Ala Glu Arg Leu His Thr Val Gly Gly Pro Gly Ser
155 160 165
Ala Arg Pro Arg Ala Phe Ser His Ser Gly Val His Ser Leu Asp
170 175 180
Gly Gly Glu Val Asp Ser Gln Ala Leu Gln Glu Leu Thr Gln Met
185 190 195
Val Ser Gly Pro Ala Ser Tyr Ser Gly Pro Lys Pro Ser Thr Gln
200 205 210
Tyr Gly Ala Pro Gly Pro Phe Ala Ala Pro Gly Glu Gly Gly Ala
215 220 225
Leu Ala Ala Thr Gly Arg Pro Pro Leu Leu Pro Thr Arg Ala Ser
230 235 240
Arg Ser Gln Arg Ala Ala Ser Glu Asp Met Thr Ser Asp Glu Glu
245 250 255
Arg Met Val Ile Cys Glu Glu Glu Gly Asp Asp Asp Val Ile Ala
260 265 270
Asp Asp Gly Phe Gly Pro Thr Asp Leu Asp Leu Lys Cys Lys Glu
275 280 285
Arg Val Thr Asp Ser Glu Ser Gly Asp Ser Ser Gly Glu Asp Pro
290 295 300
Glu Gly Asn Lys Gly Phe Gly Arg Lys Val Phe Ser Pro Val Ile
305 310 315
Arg Ser Ser Phe Thr His Cys Arg Pro Pro Leu Asp Pro Glu Pro
320 325 330
Pro Gly Pro Pro Asp Pro Pro Val Ala Phe Gly Lys Gly Tyr Gly
335 340 345
Ser Ala Pro Ser Ser Ser Ala Ser Ser Pro Ala Ser Ser Ser Ala
350 355 360
Ser Ala Ala Thr Ser Phe Ser Leu Gly Ser Gly Thr Phe Lys Ala
365 370 375
Gln Glu Ser Gly Gln Gly Ser Thr Ala Gly Pro Leu Arg Pro Pro
380 385 390
Pro Pro Gly Ala Gly Gly Pro Ala Thr Pro Ser Lys Ala Thr Arg
395 400 405
Phe Leu Pro Met Asp Pro Ala Thr Phe Arg Arg Lys Arg Pro Glu
410 415 420
Ser Val Gly Gly Leu Glu Pro Pro Gly Pro Ser Val Ile Ala Ala
425 430 435
Pro Pro Ser Gly Gly Gly Asn Ile Leu Gln Thr Leu Val Leu Pro

```
                440                 445                 450
Pro Asn Lys Glu Glu Gln Glu Gly Gly Gly Ala Arg Val Pro Ser
                455                 460                 465
Ala Pro Ala Pro Ser Leu Ala Tyr Gly Ala Pro Ala Ala Pro Leu
                470                 475                 480
Ser Arg Pro Ala Ala Thr Met Val Thr Asn Val Val Arg Pro Val
                485                 490                 495
Ser Ser Thr Pro Val Pro Ile Ala Ser Lys Pro Phe Pro Thr Ser
                500                 505                 510
Gly Arg Ala Glu Ala Ser Pro Asn Asp Thr Ala Gly Ala Arg Thr
                515                 520                 525
Glu Met Gly Thr Gly Ser Arg Val Pro Gly Gly Ser Pro Leu Gly
                530                 535                 540
Val Ser Leu Val Tyr Ser Asp Lys Lys Ser Ala Ala Ala Thr Ser
                545                 550                 555
Pro Ala Pro His Leu Val Ala Gly Pro Leu Leu Gly Thr Val Gly
                560                 565                 570
Lys Ala Pro Ala Thr Val Thr Asn Leu Leu Val Gly Thr Pro Gly
                575                 580                 585
Tyr Gly Ala Pro Ala Pro Pro Ala Val Gln Phe Ile Ala Gln Gly
                590                 595                 600
Ala Pro Gly Gly Gly Thr Thr Ala Gly Ser Gly Ala Gly Ala Gly
                605                 610                 615
Ser Gly Pro Asn Gly Pro Val Pro Leu Gly Ile Leu Gln Pro Gly
                620                 625                 630
Ala Leu Gly Lys Ala Gly Gly Ile Thr Gln Val Gln Tyr Ile Leu
                635                 640                 645
Pro Thr Leu Pro Gln Gln Leu Gln Val Ala Pro Ala Pro Ala Pro
                650                 655                 660
Ala Pro Gly Thr Lys Ala Ala Ala Pro Met Arg Pro Cys Thr His
                665                 670                 675
His Gln His Pro Phe His Pro Pro Thr Gly His Phe His Gln Arg
                680                 685                 690
Gln Ser Pro Gly Cys His Cys Thr His Ser Trp His Pro His Pro
                695                 700                 705
Ala Val Cys Thr Leu Arg Pro Thr Pro Gln Ser Pro Val Ser Phe
                710                 715                 720
Ser Arg Ala Gly Pro Ala Pro Gly Trp Leu Ser Pro Ala Ala Ala
                725                 730                 735
Trp Glu Gly Pro Ser Ala Ser Gly Arg Pro
                740                 745
```

<210> 81
<211> 256
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:334386.1.orf3:2000MAY01

<400> 81

```
Leu Ala Met Lys Asp Met Leu Thr Val Val Asp Leu Leu Leu Glu
  1                 5                  10                  15
Gly Gly Ala Asp Val Asp His Thr Asp Asn Asn Gly Arg Thr Pro
                 20                  25                  30
Leu Leu Ala Ala Ala Ser Met Gly His Ala Ser Val Val Asn Thr
                 35                  40                  45
Leu Leu Phe Trp Gly Ala Ala Val Asp Ser Ile Asp Ser Glu Gly
                 50                  55                  60
Arg Thr Val Leu Ser Ile Ala Ser Ala Gln Gly Asn Val Glu Val
                 65                  70                  75
Val Arg Thr Leu Leu Asp Arg Gly Leu Asp Glu Asn His Arg Asp
                 80                  85                  90
Asp Ala Gly Trp Thr Pro Leu His Met Ala Ala Phe Glu Gly His
                 95                 100                 105
Arg Leu Ile Cys Glu Ala Leu Ile Glu Gln Gly Ala Arg Thr Asn
```

```
                110                 115                 120
Glu Ile Asp Asn Asp Gly Arg Ile Pro Phe Ile Leu Ala Ser Gln
                125                 130                 135
Glu Gly His Tyr Asp Cys Val Gln Ile Leu Leu Glu Asn Lys Ser
                140                 145                 150
Asn Ile Asp Gln Arg Gly Tyr Asp Gly Arg Asn Ala Leu Arg Val
                155                 160                 165
Ala Ala Leu Glu Gly His Arg Asp Ile Val Glu Leu Leu Phe Ser
                170                 175                 180
His Gly Ala Asp Val Asn Cys Lys Asp Ala Asp Gly Arg Pro Thr
                185                 190                 195
Leu Tyr Ile Leu Ala Leu Glu Asn Gln Leu Thr Met Ala Glu Tyr
                200                 205                 210
Phe Leu Glu Asn Gly Ala Asn Val Glu Ala Ser Asp Ala Glu Gly
                215                 220                 225
Arg Thr Ala Leu His Val Ser Cys Trp Gln Gly His Met Gly Asn
                230                 235                 240
Gly Ala Gly Pro Asp Ser Ile Pro Cys Arg Arg Gln Cys Cys Arg
                245                 250                 255
Gln
```

<210> 82
<211> 235
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:347572.1.orf1:2000MAY01

<400> 82

```
Met Pro Ile Leu Pro Ile Ser Val Gln Leu Asp Ala Ser Leu Leu
  1               5                  10                  15
Ile Cys Leu Val Ile Cys Ala Gly Arg Phe Trp Thr Asn Leu Tyr
                20                  25                  30
Ser Leu Thr Val Pro Phe Gly Gln Lys Pro Asn Ile Asp Val Thr
                35                  40                  45
Asp Ala Met Val Asp Gln Ala Trp Asp Ala Gln Arg Ile Phe Lys
                50                  55                  60
Glu Ser Ala Glu Leu Leu Cys Ile Cys Trp Ser Ser Leu Tyr Asp
                65                  70                  75
Ser Arg Ile Leu Arg Gln Ile Pro Cys Tyr Thr Asp Pro Gly Asn
                80                  85                  90
Val Gln Lys Ala Leu Cys His Pro His Ser Leu Gly Pro Gly Glu
                95                 100                 105
Gly Arg Leu Gln Arg Ser Leu Cys Ala Gln Arg Val Thr Met Asp
                110                115                 120
Asp Phe Leu Thr Ala His His Glu Met Gly His Ile Gln Tyr Asp
                125                130                 135
Met Ala Tyr Ala Gly Gln Pro Phe Ser Ala Lys Glu Met Glu Leu
                140                145                 150
Asn Glu Gly Phe His Glu Ala Val Gly Glu Ile Met Ser Leu Ser
                155                160                 165
Ala Ala Thr Pro Lys His Leu Lys Ser Ile Gly Leu Leu Ser Pro
                170                175                 180
Glu Phe Ser Thr Asn Asp Asn Glu Thr Glu Ile Asn Phe Leu Leu
                185                190                 195
Lys Gln Ala Leu Thr Ile Val Gly Thr Leu Pro Phe Thr Tyr Met
                200                205                 210
Leu Glu Lys Trp Arg Trp Met Val Phe Lys Arg Gly Asn Ser Gln
                215                220                 225
Arg Pro Val Gly Glu Lys Gly Gly Gly Arg
                230                235
```

<210> 83
<211> 617

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:817314.1.orf1:2000MAY01

<400> 83
Asn Met Ala Gln Phe Tyr Tyr Lys Arg Asn Val Asn Ala Pro Tyr
1 5 10 15
Arg Asp Arg Ile Pro Leu Arg Ile Val Arg Ala Glu Ser Glu Leu
20 25 30
Ser Pro Ser Glu Lys Ala Tyr Leu Asn Ala Val Glu Lys Gly Asp
35 40 45
Tyr Ala Ser Val Lys Lys Ser Leu Glu Glu Ala Glu Ile Tyr Phe
50 55 60
Lys Ile Asn Ile Asn Cys Ile Asp Pro Leu Gly Arg Thr Ala Leu
65 70 75
Leu Ile Ala Ile Glu Asn Glu Asn Leu Glu Leu Ile Glu Leu Leu
80 85 90
Leu Ser Phe Asn Val Tyr Val Gly Asp Ala Leu Leu His Ala Ile
95 100 105
Arg Lys Glu Val Val Gly Ala Val Glu Leu Leu Leu Asn His Lys
110 115 120
Lys Pro Ser Gly Glu Lys Gln Val Pro Pro Ile Leu Leu Asp Lys
125 130 135
Gln Phe Ser Glu Phe Thr Pro Asp Ile Thr Pro Ile Ile Leu Ala
140 145 150
Ala His Thr Asn Asn Tyr Glu Ile Ile Lys Leu Leu Val Gln Lys
155 160 165
Gly Val Ser Val Pro Arg Pro His Glu Val Arg Cys Asn Cys Val
170 175 180
Glu Cys Val Ser Ser Ser Asp Val Asp Ser Leu Arg His Ser Arg
185 190 195
Ser Arg Leu Asn Ile Tyr Lys Ala Leu Ala Ser Pro Ser Leu Ile
200 205 210
Ala Leu Ser Ser Glu Asp Pro Phe Leu Thr Ala Phe Gln Leu Ser
215 220 225
Trp Glu Leu Gln Glu Leu Ser Lys Val Glu Asn Glu Phe Lys Ser
230 235 240
Glu Tyr Glu Glu Leu Ser Arg Gln Cys Lys Gln Phe Ala Lys Asp
245 250 255
Leu Leu Asp Gln Thr Arg Ser Ser Arg Glu Leu Glu Ile Ile Leu
260 265 270
Asn Tyr Arg Asp Asp Asn Ser Leu Ile Glu Glu Gln Ser Gly Asn
275 280 285
Asp Leu Ala Arg Leu Lys Leu Ala Ile Lys Tyr Arg Gln Lys Glu
290 295 300
Phe Val Ala Gln Pro Asn Cys Gln Gln Leu Leu Ala Ser Arg Trp
305 310 315
Tyr Asp Glu Phe Pro Gly Trp Arg Arg Arg His Trp Ala Val Lys
320 325 330
Met Val Thr Cys Phe Ile Ile Gly Leu Leu Phe Pro Val Phe Ser
335 340 345
Val Cys Tyr Leu Ile Ala Pro Lys Ser Pro Leu Gly Leu Phe Ile
350 355 360
Arg Lys Pro Phe Ile Lys Phe Ile Cys His Thr Ala Ser Tyr Leu
365 370 375
Thr Phe Leu Phe Leu Leu Leu Leu Ala Ser Gln His Ile Asp Arg
380 385 390
Ser Asp Leu Asn Arg Gln Gly Pro Pro Pro Thr Ile Val Glu Trp
395 400 405
Met Ile Leu Pro Trp Val Leu Gly Phe Ile Trp Gly Glu Ile Lys
410 415 420
Gln Met Trp Asp Gly Gly Leu Gln Asp Tyr Ile His Asp Trp Trp
425 430 435
Asn Leu Met Asp Phe Val Met Asn Ser Leu Tyr Leu Ala Thr Ile

```
                440                 445                 450
Ser Leu Lys Ile Val Ala Phe Val Lys Tyr Ser Ala Leu Asn Pro
                455                 460                 465
Arg Glu Ser Trp Asp Met Trp His Pro Thr Leu Val Ala Glu Ala
                470                 475                 480
Leu Phe Ala Ile Ala Asn Ile Phe Ser Ser Leu Arg Leu Ile Ser
                485                 490                 495
Leu Phe Thr Ala Asn Ser His Leu Gly Pro Leu Gln Ile Ser Leu
                500                 505                 510
Gly Arg Met Leu Leu Asp Ile Leu Lys Phe Leu Phe Ile Tyr Cys
                515                 520                 525
Leu Val Leu Leu Ala Phe Ala Asn Gly Leu Asn Gln Leu Tyr Phe
                530                 535                 540
Tyr Tyr Glu Glu Thr Lys Gly Leu Thr Cys Lys Gly Ile Arg Cys
                545                 550                 555
Glu Lys Gln Asn Asn Ala Phe Ser Thr Leu Phe Glu Thr Leu Gln
                560                 565                 570
Ser Leu Phe Trp Ser Ile Phe Gly Leu Ile Asn Leu Tyr Val Thr
                575                 580                 585
Asn Val Lys Ala Gln His Glu Phe Thr Glu Phe Val Gly Ala Thr
                590                 595                 600
Leu Phe Gly Asp Ile Thr Met Ser Ser Leu Trp Leu Phe Tyr Ser
                605                 610                 615
Thr Cys
```

<210> 84
<211> 293
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:000290.1.orf3:2000MAY01

<400> 84

```
Gly Ala His Ala Lys Thr Gly Ile Gln Ile Gly Met Leu Ser Thr
  1                 5                  10                  15
Gly Lys Asp Arg Ser Leu Arg Val Thr Gly Met Thr Trp Arg Ser
                 20                  25                  30
Ser Tyr Val Pro Val Ser Ala Pro Pro Pro Asn Ser Ser Glu Gln
                 35                  40                  45
Tyr Ser Ser Gly Ala Gln Ser Ile Pro Ser Thr Val Thr Val Ile
                 50                  55                  60
Ala Pro Trp Ser Pro Thr Leu Glu Asn Thr Thr Trp Glu Leu Val
                 65                  70                  75
Leu Leu Leu Leu Lys Ile Ile Ser Ser Ser Asn Ser Phe Gly Arg
                 80                  85                  90
Asn Leu Pro Pro Lys Arg Arg Cys Arg Asp Tyr Asp Glu Arg Gly
                 95                 100                 105
Phe Cys Val Leu Gly Asp Leu Cys Gln Phe Asp His Gly Asn Asp
                110                 115                 120
Pro Leu Val Val Asp Glu Val Ala Leu Pro Ser Met Ile Pro Phe
                125                 130                 135
Pro Pro Pro Pro Pro Gly Leu Pro Pro Pro Thr Thr Pro Gly Met
                140                 145                 150
Leu Met Pro Pro Met Pro Gly Pro Gly Pro Gly Pro Gly Pro Gly
                155                 160                 165
Pro Gly Pro Gly Pro Gly Pro Gly Pro Gly Pro Gly His Ser Met
                170                 175                 180
Arg Leu Pro Val Pro Gln Gly His Gly Gln Pro Pro Pro Ser Val
                185                 190                 195
Val Leu Pro Ile Pro Arg Pro Pro Ile Thr Gln Ser Ser Leu Ile
                200                 205                 210
Asn Ser Arg Asp Gln Pro Gly Thr Ser Ala Val Pro Asn Leu Ala
                215                 220                 225
Ser Val Gly Thr Arg Leu Pro Pro Pro Leu Pro Gln Asn Leu Leu
```

```
                230                 235                 240
Tyr Thr Val Ser Glu Arg Gln Pro Met Tyr Ser Arg Glu His Gly
                245                 250                 255
Ala Ala Ala Ser Glu Arg Leu Gln Leu Gly Thr Pro Pro Pro Leu
                260                 265                 270
Leu Ala Ala Arg Leu Val Pro Pro Arg Asn Leu Met Gly Ser Ser
                275                 280                 285
Ile Gly Tyr His Thr Ser Val Ser
                290
```

<210> 85
<211> 276
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:023518.3.orf3:2000MAY01

```
<400> 85
Leu Ser Pro Asp Arg Leu Leu Val Leu Pro Asp Asn Tyr Ser His
  1                 5                  10                  15
Phe Ser Gln Ala Ser Ala Asn Leu Gln Gly Pro Ser Arg Thr Thr
                 20                  25                  30
Glu Leu Phe His Pro Thr Leu Ala Ser Ile Ser Ser Pro Met Leu
                 35                  40                  45
Glu Gly Ala Glu Leu Tyr Phe Asn Val Asp His Gly Tyr Leu Glu
                 50                  55                  60
Gly Leu Val Arg Gly Cys Lys Ala Ser Leu Leu Thr Gln Gln Asp
                 65                  70                  75
Tyr Ile Asn Leu Val Gln Cys Glu Thr Leu Glu Ala Pro Phe Phe
                 80                  85                  90
Gln Asp Cys Met Ser Glu Asn Ala Leu Asp Glu Leu Asn Ile Glu
                 95                 100                 105
Leu Leu Arg Asn Lys Leu Tyr Lys Ser Tyr Leu Glu Ala Phe Tyr
                110                 115                 120
Lys Phe Cys Lys Asn His Gly Asp Val Thr Ala Glu Val Met Cys
                125                 130                 135
Pro Ile Leu Glu Phe Glu Ala Asp Arg Arg Ala Phe Ile Ile Thr
                140                 145                 150
Leu Asn Ser Phe Gly Thr Glu Leu Ser Lys Glu Asp Arg Glu Thr
                155                 160                 165
Leu Tyr Pro Thr Phe Arg Gln Leu Tyr Pro Glu Gly Leu Arg Leu
                170                 175                 180
Leu Ala Gln Ala Glu Asp Phe Asp Gln Met Lys Asn Val Ala Asp
                185                 190                 195
His Tyr Gly Val Tyr Lys Pro Leu Phe Glu Ala Val Gly Gly Ser
                200                 205                 210
Gly Gly Lys Thr Leu Glu Asp Val Phe Tyr Glu Arg Glu Val Gln
                215                 220                 225
Met Asn Val Leu Ala Phe Asn Arg Gln Phe His Tyr Gly Val Phe
                230                 235                 240
Tyr Ala Tyr Val Lys Leu Lys Glu Gln Glu Ile Arg Asn Ile Val
                245                 250                 255
Trp Ile Ala Glu Cys Ile Ser Gln Arg His Arg Thr Lys Ile Asn
                260                 265                 270
Ser Tyr Ile Pro Ile Leu
                275
```

<210> 86
<211> 355
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:1084246.1.orf3:2000MAY01

<400> 86
Pro Leu Asp Arg Glu Thr Ser Thr Glu Tyr Asn Ile Thr Ile Ala
1 5 10 15
Val Thr Asp Leu Gly Thr Pro Arg Leu Lys Thr Gln Gln Asn Ile
20 25 30
Thr Val Gln Val Ser Asp Val Asn Asp Asn Ala Pro Ala Phe Thr
35 40 45
Gln Thr Ser Tyr Thr Leu Phe Val Arg Glu Asn Asn Ser Pro Ala
50 55 60
Leu His Ile Gly Ser Val Ser Ala Thr Asp Arg Asp Ser Gly Thr
65 70 75
Asn Ala Gln Val Thr Tyr Ser Leu Leu Pro Pro Gln Asp Pro His
80 85 90
Leu Pro Leu Ala Ser Leu Val Ser Ile Asn Ala Asp Asn Gly His
95 100 105
Leu Phe Ala Leu Arg Ser Leu Asp Tyr Glu Ala Leu Gln Ala Phe
110 115 120
Glu Phe Arg Val Gly Ala Ser Asp Arg Gly Ser Pro Ala Leu Ser
125 130 135
Ser Glu Ala Leu Val Arg Val Leu Val Leu Asp Thr Asn Asp Asn
140 145 150
Ser Pro Phe Val Leu Tyr Pro Leu Gln Asn Gly Ser Ala Pro Cys
155 160 165
Thr Glu Leu Val Pro Arg Ala Ala Glu Pro Gly Tyr Leu Val Thr
170 175 180
Lys Val Val Ala Val Asp Gly Asp Ser Gly Gln Asn Ala Trp Leu
185 190 195
Ser Tyr Gln Leu Leu Lys Ala Thr Glu Pro Gly Leu Phe Gly Val
200 205 210
Trp Ala His Asn Gly Glu Val Arg Thr Ala Arg Leu Leu Ser Glu
215 220 225
Arg Asp Ala Ala Lys His Arg Leu Val Val Leu Val Lys Asp Asn
230 235 240
Gly Glu Pro Pro Arg Ser Ala Thr Ala Thr Leu His Val Leu Leu
245 250 255
Val Asp Gly Phe Ser Gln Pro Tyr Leu Pro Leu Pro Glu Ala Ala
260 265 270
Pro Ala Gln Ala Gln Ala Asp Ser Leu Thr Val Tyr Leu Val Val
275 280 285
Ala Leu Ala Ser Val Ser Ser Leu Phe Leu Phe Ser Val Leu Leu
290 295 300
Phe Val Ala Val Arg Leu Cys Arg Arg Ser Arg Ala Ala Ser Val
305 310 315
Gly Arg Cys Ser Val Pro Glu Gly Pro Phe Pro Gly His Leu Val
320 325 330
Asp Val Ser Gly Thr Gly Thr Leu Ser Gln Glu Leu Pro Val Arg
335 340 345
Gly Val Ser Asp Arg Arg Leu Trp Asp Trp
350 355

<210> 87
<211> 745
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:1165828.1.orf2:2000MAY01

<400> 87
Val Phe Glu Met Leu Tyr Ser Ser Arg Gly Asp Pro Glu Gly Gln
1 5 10 15
Pro Leu Leu Leu Ser Leu Leu Ile Leu Ala Met Trp Val Val Gly
20 25 30
Ser Gly Gln Leu His Tyr Ser Val Pro Glu Glu Ala Glu His Gly
35 40 45
Thr Phe Val Gly Arg Ile Ala Gln Asp Leu Gly Leu Glu Leu Ala

```
                50                  55                  60
Glu Leu Val Pro Arg Leu Phe Gln Leu Asp Ser Lys Gly Arg Gly
                65                  70                  75
Asp Leu Leu Glu Val Asn Leu Gln Asn Gly Ile Leu Phe Val Asn
                80                  85                  90
Ser Arg Ile Asp Arg Glu Glu Leu Cys Gly Arg Ser Ala Glu Cys
                95                 100                 105
Ser Ile His Leu Glu Val Ile Val Asp Arg Pro Leu Gln Val Phe
               110                 115                 120
His Val Asp Val Glu Val Lys Asp Ile Asn Asp Asn Pro Pro Val
               125                 130                 135
Phe Pro Ala Thr Gln Lys Asn Leu Phe Ile Ala Glu Ser Arg Pro
               140                 145                 150
Leu Asp Ser Arg Phe Pro Leu Glu Gly Ala Ser Asp Ala Asp Ile
               155                 160                 165
Gly Glu Asn Ala Leu Leu Thr Tyr Arg Leu Ser Pro Asn Glu Tyr
               170                 175                 180
Phe Phe Leu Asp Val Pro Thr Ser Asn Gln Gln Val Lys Pro Leu
               185                 190                 195
Gly Leu Val Leu Arg Lys Leu Leu Asp Arg Glu Glu Thr Pro Glu
               200                 205                 210
Leu His Leu Leu Leu Thr Ala Thr Asp Gly Gly Lys Pro Glu Leu
               215                 220                 225
Thr Gly Thr Val Gln Leu Leu Ile Thr Val Leu Asp Asn Asn Asp
               230                 235                 240
Asn Ala Pro Val Phe Asp Arg Thr Leu Tyr Thr Val Lys Leu Pro
               245                 250                 255
Glu Asn Val Ser Ile Gly Thr Leu Val Ile His Pro Asn Ala Ser
               260                 265                 270
Asp Leu Asp Glu Gly Leu Asn Gly Asp Ile Ile Tyr Ser Phe Ser
               275                 280                 285
Ser Asp Val Ser Pro Asp Ile Lys Ser Lys Phe His Met Asp Pro
               290                 295                 300
Leu Ser Gly Ala Ile Thr Val Ile Gly His Met Asp Phe Glu Glu
               305                 310                 315
Ser Arg Ala His Lys Ile Pro Val Glu Ala Val Asp Lys Gly Phe
               320                 325                 330
Pro Pro Leu Ala Gly His Cys Thr Leu Leu Val Glu Val Val Asp
               335                 340                 345
Val Asn Asp Asn Ala Pro Gln Leu Thr Ile Lys Thr Leu Ser Val
               350                 355                 360
Pro Val Lys Glu Asp Ala Gln Leu Gly Thr Val Ile Ala Leu Ile
               365                 370                 375
Ser Val Ile Asp Leu Asp Ala Asp Ala Asn Gly Gln Val Thr Cys
               380                 385                 390
Ser Leu Thr Pro His Val Pro Phe Lys Leu Val Ser Thr Tyr Lys
               395                 400                 405
Asn Tyr Tyr Ser Leu Val Leu Asp Arg Ala Leu Asp Arg Glu Ser
               410                 415                 420
Val Ser Ala Tyr Glu Leu Val Val Thr Ala Arg Asp Gly Gly Ser
               425                 430                 435
Pro Ser Leu Trp Ala Thr Ala Arg Val Ser Val Glu Val Ala Asp
               440                 445                 450
Val Asn Asp Asn Ala Pro Ala Phe Ala Gln Ser Glu Tyr Thr Val
               455                 460                 465
Phe Val Lys Glu Asn Asn Pro Pro Gly Cys His Ile Phe Thr Val
               470                 475                 480
Ser Ala Arg Asp Ala Asp Ala Gln Glu Asn Ala Leu Val Ser Tyr
               485                 490                 495
Ser Leu Val Glu Arg Arg Leu Gly Glu Arg Ser Leu Ser Ser Tyr
               500                 505                 510
Val Ser Val His Ala Glu Ser Gly Lys Val Tyr Ala Leu Gln Pro
               515                 520                 525
Leu Asp His Glu Glu Leu Glu Leu Leu Gln Phe Gln Val Ser Ala
               530                 535                 540
Arg Asp Ala Gly Val Pro Pro Leu Gly Ser Asn Val Thr Leu Gln
               545                 550                 555
```

```
Val Phe Val Leu Asp Glu Asn Asp Asn Ala Pro Ala Leu Leu Thr
                560                 565                 570
Pro Arg Met Arg Gly Thr Asp Gly Ala Val Ser Glu Met Val Leu
                575                 580                 585
Arg Ser Val Gly Ala Gly Val Val Val Gly Lys Val Arg Ala Val
                590                 595                 600
Asp Ala Asp Ser Gly Tyr Asn Ala Trp Leu Ser Tyr Glu Leu Gln
                605                 610                 615
Pro Glu Thr Ala Ser Ala Ser Ile Pro Phe Arg Val Gly Leu Tyr
                620                 625                 630
Thr Gly Glu Ile Ser Thr Thr Arg Ala Leu Asp Glu Thr Asp Ala
                635                 640                 645
Pro Arg Gln Arg Leu Leu Val Leu Val Lys Asp His Gly Glu Pro
                650                 655                 660
Ala Leu Thr Ala Thr Ala Thr Val Leu Val Ser Leu Val Glu Ser
                665                 670                 675
Gly Gln Ala Pro Lys Ser Ser Ser Arg Ala Ser Val Gly Ala Thr
                680                 685                 690
Gly Pro Glu Val Thr Leu Val Asp Val Asn Val Tyr Leu Ile Ile
                695                 700                 705
Ala Ile Cys Ala Val Ser Ser Leu Leu Val Leu Thr Leu Leu Leu
                710                 715                 720
Tyr Thr Val Leu Arg Cys Ser Ala Met Pro Thr Glu Gly Glu Cys
                725                 730                 735
Ala Pro Gly Lys Ala Asp Ala Gly Val Phe
                740                 745
```

<210> 88
<211> 781
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:007302.1.orf2:2000MAY01

<400> 88

```
Asp Ser His Cys Asn Ile Met Thr Lys Asp Lys Glu Pro Ile Val
  1                 5                  10                  15
Lys Ser Phe His Phe Val Cys Leu Met Ile Ile Ile Val Gly Thr
                 20                  25                  30
Arg Ile Gln Phe Ser Asp Gly Asn Glu Phe Ala Val Asp Lys Ser
                 35                  40                  45
Lys Arg Gly Leu Ile His Val Pro Lys Asp Leu Pro Leu Lys Thr
                 50                  55                  60
Lys Val Leu Asp Met Ser Gln Asn Tyr Ile Ala Glu Leu Gln Val
                 65                  70                  75
Ser Asp Met Ser Phe Leu Ser Glu Leu Thr Val Leu Arg Leu Ser
                 80                  85                  90
His Asn Arg Ile Gln Leu Leu Asp Leu Ser Val Phe Lys Phe Asn
                 95                 100                 105
Gln Asp Leu Glu Tyr Leu Asp Leu Ser His Asn Gln Leu Gln Lys
                110                 115                 120
Ile Ser Cys His Pro Ile Val Ser Phe Arg His Leu Asp Leu Ser
                125                 130                 135
Phe Asn Asp Phe Lys Ala Leu Pro Ile Cys Lys Glu Phe Gly Asn
                140                 145                 150
Leu Ser Gln Leu Asn Phe Leu Gly Leu Ser Ala Met Lys Leu Gln
                155                 160                 165
Lys Leu Asp Leu Leu Pro Ile Ala His Leu His Leu Ser Tyr Ile
                170                 175                 180
Leu Leu Asp Leu Arg Asn Tyr Tyr Ile Lys Glu Asn Glu Thr Glu
                185                 190                 195
Ser Leu Gln Ile Leu Asn Ala Lys Thr Leu His Leu Val Phe His
                200                 205                 210
Pro Thr Ser Leu Phe Ala Ile Gln Val Asn Ile Ser Val Asn Thr
                215                 220                 225
```

```
Leu Gly Cys Leu Gln Leu Thr Asn Ile Lys Leu Asn Asp Asp Asn
                230                 235                 240
Cys Gln Val Phe Ile Lys Phe Leu Ser Glu Leu Thr Arg Gly Pro
                245                 250                 255
Thr Leu Leu Asn Phe Thr Leu Asn His Ile Glu Thr Thr Trp Lys
                260                 265                 270
Cys Leu Val Arg Val Phe Gln Phe Leu Trp Pro Lys Pro Val Glu
                275                 280                 285
Tyr Leu Asn Ile Tyr Asn Leu Thr Ile Ile Glu Ser Ile Arg Glu
                290                 295                 300
Glu Asp Phe Thr Tyr Ser Lys Thr Thr Leu Lys Ala Leu Thr Ile
                305                 310                 315
Glu His Ile Thr Asn Gln Val Phe Leu Phe Ser Gln Thr Ala Leu
                320                 325                 330
Tyr Thr Val Phe Ser Glu Met Asn Ile Met Met Leu Thr Ile Ser
                335                 340                 345
Asp Thr Pro Phe Ile His Met Leu Cys Pro His Ala Pro Ser Thr
                350                 355                 360
Phe Lys Phe Leu Asn Phe Thr Gln Asn Val Phe Thr Asp Ser Ile
                365                 370                 375
Phe Glu Lys Cys Ser Thr Leu Val Lys Leu Glu Thr Leu Ile Leu
                380                 385                 390
Gln Lys Asn Gly Leu Lys Asp Leu Phe Lys Val Gly Leu Met Thr
                395                 400                 405
Lys Asp Met Pro Ser Leu Glu Ile Leu Asp Val Ser Trp Asn Ser
                410                 415                 420
Leu Glu Ser Gly Arg His Lys Glu Asn Cys Thr Trp Val Glu Ser
                425                 430                 435
Ile Val Val Leu Asn Leu Ser Ser Asn Met Leu Thr Asp Ser Val
                440                 445                 450
Phe Arg Cys Leu Pro Pro Arg Ile Lys Val Leu Asp Leu His Ser
                455                 460                 465
Asn Lys Ile Lys Ser Val Pro Lys Gln Val Val Lys Leu Glu Ala
                470                 475                 480
Leu Gln Glu Leu Asn Val Ala Phe Asn Ser Leu Thr Asp Leu Pro
                485                 490                 495
Gly Cys Gly Ser Phe Ser Ser Leu Ser Val Leu Ile Ile Asp His
                500                 505                 510
Asn Ser Val Ser His Pro Ser Ala Asp Phe Phe Gln Ser Cys Gln
                515                 520                 525
Lys Met Arg Ser Ile Lys Ala Gly Asp Asn Pro Phe Gln Cys Thr
                530                 535                 540
Cys Glu Leu Arg Glu Phe Val Lys Asn Ile Asp Gln Val Ser Ser
                545                 550                 555
Glu Val Leu Glu Gly Trp Pro Asp Ser Tyr Lys Cys Asp Tyr Pro
                560                 565                 570
Glu Ser Tyr Arg Gly Ser Pro Leu Lys Asp Phe His Met Ser Glu
                575                 580                 585
Leu Ser Cys Asn Ile Thr Leu Leu Ile Val Thr Ile Gly Ala Thr
                590                 595                 600
Met Leu Val Leu Ala Val Thr Val Thr Ser Leu Cys Ile Tyr Leu
                605                 610                 615
Asp Leu Pro Trp Tyr Leu Arg Met Val Cys Gln Trp Thr Gln Thr
                620                 625                 630
Arg Arg Arg Ala Arg Asn Ile Pro Leu Glu Glu Leu Gln Arg Asn
                635                 640                 645
Leu Gln Phe His Ala Phe Ile Ser Tyr Ser Glu His Asp Ser Ala
                650                 655                 660
Trp Val Lys Ser Glu Leu Val Pro Tyr Leu Glu Lys Glu Asp Ile
                665                 670                 675
Gln Ile Cys Leu His Glu Arg Asn Phe Val Pro Gly Lys Ser Ile
                680                 685                 690
Val Glu Asn Ile Ile Asn Cys Ile Glu Lys Ser Tyr Lys Ser Ile
                695                 700                 705
Phe Val Leu Ser Pro Asn Phe Val Gln Ser Glu Trp Cys His Tyr
                710                 715                 720
Glu Leu Tyr Phe Ala His His Asn Leu Phe His Glu Gly Ser Asn
```

```
                725                 730                 735
Asn Leu Ile Leu Ile Leu Leu Glu Pro Ile Pro Gln Asn Ser Ile
                740                 745                 750
Pro Asn Lys Tyr His Lys Leu Lys Ala Leu Met Thr Gln Arg Thr
                755                 760                 765
Tyr Leu Gln Trp Pro Lys Glu Lys Ser Lys Arg Gly Ala Leu Leu
                770                 775                 780
Gly
```

<210> 89
<211> 293
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: LI:236386.4.orf2:2000MAY01

<400> 89

```
Trp His Lys Ile Ala Glu Thr Tyr Ser Ile Glu Met Gly Pro Arg
  1                 5                  10                  15
Gly Pro Gln Cys Glu Gly Ala Ile Pro Thr His Leu Pro Ala Leu
                 20                  25                  30
Trp Arg Thr Pro Gln Asn Arg Pro Asn Ser Arg Ala Ser Lys Ala
                 35                  40                  45
Thr Ser Pro Thr Ser Ser His Pro Pro Met Leu Pro His Pro Ser
                 50                  55                  60
Thr Gly Ala Thr Asn Thr Leu Thr Gly Ser Ile Thr Arg Leu Leu
                 65                  70                  75
His Lys Phe Thr Val Ile Ser Val Pro His Leu Pro Glu Lys Gln
                 80                  85                  90
Ala Thr Gly Arg Phe Glu Glu Asp Phe Ile Glu Lys Arg Lys Arg
                 95                 100                 105
Arg Leu Ile Leu Trp Met Asp His Met Thr Ser His Pro Val Leu
                110                 115                 120
Ser Gln Tyr Glu Gly Phe Gln His Phe Leu Ser Cys Leu Asp Asp
                125                 130                 135
Lys Gln Trp Lys Met Gly Lys Arg Arg Ala Glu Lys Asp Glu Met
                140                 145                 150
Val Gly Ala Ser Phe Leu Leu Thr Phe Gln Ile Pro Thr Glu His
                155                 160                 165
Gln Asp Leu Gln Asp Val Glu Asp Arg Val Asp Thr Phe Lys Ala
                170                 175                 180
Phe Ser Lys Lys Met Asp Asp Ser Val Leu Gln Leu Ser Thr Val
                185                 190                 195
Ala Ser Glu Leu Val Arg Lys His Val Gly Gly Phe Pro Gln Gly
                200                 205                 210
Ile Pro Glu Arg Trp Ala Val Pro Ser Arg Pro Ser Val Ile Pro
                215                 220                 225
Ser Arg Trp Thr Pro Pro Phe Ala Leu Arg Pro Ser Thr Val Pro
                230                 235                 240
Phe Leu Thr Arg Ala Val Pro Met Lys Pro Ser Gly Arg Cys Leu
                245                 250                 255
Leu Ser Ser Pro Arg Met Thr Ser Ser Arg Cys Trp Thr His Cys
                260                 265                 270
Leu Ser Thr Arg Ala Cys Ser Pro Thr Ser Leu Thr Ser Ser Ile
                275                 280                 285
Tyr Lys Lys Ala Pro Ser Pro Arg
                290
```

<210> 90
<211> 241
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: LI:252904.5.orf1:2000MAY01

<400> 90

```
Thr Pro Cys Leu Gln Glu Val Ala Gly Pro Leu Leu Asp Gly Met
  1               5                  10                  15
Val Tyr Ala Leu Gly Gly Met Gly Pro Asp Thr Ala Pro Gln Ala
             20                  25                  30
Gln Val Arg Val Tyr Glu Pro Arg Arg Asp Cys Trp Leu Ser Leu
             35                  40                  45
Pro Ser Met Pro Thr Pro Cys Tyr Gly Ala Ser Thr Phe Leu His
             50                  55                  60
Gly Asn Lys Ile Tyr Val Leu Gly Gly Arg Gln Gly Lys Leu Pro
             65                  70                  75
Val Thr Ala Phe Glu Ala Phe Asp Leu Glu Ala Arg Thr Trp Thr
             80                  85                  90
Arg His Pro Ser Leu Pro Ser Arg Arg Ala Phe Ala Gly Cys Ala
             95                 100                 105
Met Ala Glu Gly Ser Val Phe Ser Leu Gly Gly Leu Gln Gln Pro
            110                 115                 120
Gly Pro His Asn Phe Tyr Ser Arg Pro His Phe Val Asn Thr Val
            125                 130                 135
Glu Met Phe Asp Leu Glu His Gly Ser Trp Thr Lys Leu Pro Arg
            140                 145                 150
Ser Leu Arg Met Arg Asp Lys Arg Ala Asp Phe Val Val Gly Ser
            155                 160                 165
Leu Gly Gly His Ile Val Ala Ile Gly Gly Leu Gly Asn Gln Pro
            170                 175                 180
Cys Pro Leu Gly Ser Val Glu Ser Phe Ser Leu Ala Arg Arg Arg
            185                 190                 195
Trp Glu Ala Leu Pro Ala Met Pro Thr Ala Arg Cys Ser Cys Ser
            200                 205                 210
Ser Leu Gln Ala Gly Pro Arg Leu Phe Val Ile Gly Gly Val Ala
            215                 220                 225
Gln Gly Pro Ser Gln Ala Val Glu Ala Leu Cys Leu Arg Asp Gly
            230                 235                 240
Val
```

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/24228

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :A61K 39/00, 39/02; C12P 21/00, 1/21

US CL : 424/184.1, 190.1, 192.1; 435/69.1, 252.3; 536/23.5

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/184.1, 190.1, 192.1; 435/69.1, 252.3; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, BIOSIS, EMBASE, SCISERACH, CAPLUS ON STN

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,601,831 A (GREEN et al) 11 February 1997 (2/11/1997), see entire document.	1-11, 16-18, 22-26, 29-31, 37-40, 43-44, 50-52, 56-58 64-69, 80 and 82-84
X	EP 0,540,128 A1 (BIOTECHNOLOGY AUSTRALIA PTY. LTD.) 05 May 1993 (05/05/93), see entire document, page 20, lines 26-48, in particular.	1-12, 16-18, 22-31, 37-40, 50 64-72, 76-78, 80 and 82-84

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

27 SEPTEMBER 2001

Date of mailing of the international search report

16 NOV 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

PHUONG N. HUYNH

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)*

INTERNATIONAL SEARCH REPORT

Internatic...l application No.
PCT/US01/24228

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97/06590 A1 (BIOENTERPRISES PTY. LTD.) 05 November 1987 (05.11.1987), see entire document, page 16, see claims 1-35, claims 38-42, in particular.	1-12, 16-18, 22-31, 37-40, 50, 55, 59-61, 64-72, 75-78 and 82-84
Y	NAKAMURA, K et al. DNA Sequence of the Gene for the Outer Membrane Lipoprotein of E. Coli an Extremely AT-Rich Promoter. Cell. December 1979, Vol. 18, pages 1109-1117, see page 1114, in particular.	13, 19-21 and 62
Y	MEEKER A et al. A Fusion Protein Between Serum Amyloid A and Staphylococcal Nuclease - Synthesis, Purification, and Structural Studies. Proteins. March 1998, Vol. 30 No. 4, pages 381-7, see entire document.	14, 34, 41-42, 49, 73 and 79
Y	VERMA, N et al. Delivery of class I and class II MHC-restricted T-cell epitopes of listeriolysin of Listeria monocytogenes by attenuated salmonella. Vaccine. 1995, Vol. 13, No. 2, pages 142-150, see entire document.	14-15, 34-36, 41-42, 49, 63 and 73
Y	US 5,693,495 A (BREITENEDER et al) 02 December 1997 (2.12.1997), see entire document.	32
Y	US 5,877,289 A (THORPE et al) 02 March 1999 (2.3.1999), see entire document.	33, 45-49, 53-54 and 73-74